University of Namur

MASTER EN SCIENCES MATHÉMATIQUES

Méthodes de linéarisation pour résoudre un problème d'équilibre mathématique non différentiable

Nélisse, Stéphanie

Award date:
2007

FUNDP
NAMUR

Faculté des Sciences
Département de Mathématique

Rempart de la Vierge, 8
B - 5000 Namur (Belgique)

Méthodes de linéarisation pour résoudre un problème d'équilibre mathématique non différentiable

Mémoire présenté pour l'obtention
du grade de
Licencié en Sciences Mathématiques
par

NÉLISSE **Stéphanie**

Promoteur : STRODIOT Jean-Jacques

Année Académique 2006-2007

Résumé

Nous présentons deux méthodes de linéarisation pour résoudre un problème d'équilibre mathématique non différentiable de la forme : trouver x^* dans un ensemble C tel que $f(x^*, y) \geq 0$ pour tout y dans C. Ce problème recouvre de nombreux problèmes dont les problèmes d'inéquations variationnelles. La première méthode utilisée est la méthode de relaxation combinée. Nous décrivons cette méthode pour le problème d'équilibre, les problèmes d'inéquations variationnelles généralisés et multivoques ainsi que les inclusions multivoques. La seconde méthode utilisée est la méthode faisceau. Nous décrivons tout d'abord cette méthode pour le problème d'équilibre puis nous l'appliquons aux problèmes d'inéquations variationnelles généralisés et multivoques.

Mots clés : problème d'équilibre, problème d'inéquation variationnelle généralisé, problème d'inéquation variationnelle multivoque, inclusions multivoques, méthode de relaxation combinée, méthode faisceau.

Abstract

We present two linearization methods to solve a nondifferentiable equilibrium problem of the form : find x^* in a set C such that $f(x^*, y) \geq 0$ for all y in C. This problem covers many problems as the variational inequalities problems. The first method is the combined relaxation method. We describe this method for the equilibrium problem, the generalized and multivalued variational inequalities problems and the multivalued inclusions. The second method is the bundle method. We describe it first for the equilibrium problem then we apply it to the generalized and multivalued variational inequalities problems.

Keywords : equilibrium problem, generalized variational inequality problem, multivalued variational inequality problem, multivalued inclusions, combined relaxation method, bundle method.

Remerciements

Nous remercions tout particulièrement Monsieur Strodiot, notre promoteur, pour son aide tout au long du travail. Ses conseils avisés et ses explications nous ont été très précieux.

Nous remercions également nos parents pour leur soutien.

Table des matières

Introduction

Nous proposons de résoudre un problème d'équilibre mathématique de la forme : trouver x^* dans un ensemble C tel que $f(x^*, y) \geq 0$ pour tout y dans C. Ce problème recouvre de nombreux problèmes connus dont les problèmes d'inéquations variationnelles.

Deux méthodes sont utilisées pour résoudre un tel problème : la méthode de relaxation combinée et la méthode faisceau.

Dans le premier chapitre, nous définissons le problème d'équilibre ainsi que quelques problèmes apparentés.

Au second chapitre, nous décrivons un schéma de base pour construire une méthode de relaxation combinée afin de trouver des points d'équilibre sujets à des contraintes linéaires. Nous étudions ensuite la convergence de cette méthode. Le schéma de base est ensuite modifié afin d'obtenir une méthode combinée de non-descente pour trouver des points d'équilibre. Cette modification permet d'accélérer la convergence de la méthode.

Dans le troisième chapitre, nous décrivons une méthode de relaxation combinée pour le problème d'inéquation variationnelle généralisé. Nous étudions la convergence de la méthode ainsi que le taux de convergence.

Au chapitre quatre, nous décrivons une méthode de relaxation combinée mais cette fois pour le problème d'inéquation variationnelle multivoque. Nous établissons quelques propriétés des opérateurs auxiliaires que nous utiliserons pour prouver la convergence de la méthode. Nous donnons ensuite certaines estimations de la complexité de la méthode que nous avons établie. Enfin, nous modifions la méthode pour résoudre le problème d'inéquation variationnelle multivoque.

Dans le cinquième chapitre, nous décrivons une méthode de relaxation combinée inexacte pour les inclusions multivoques et nous en étudions la convergence.

Dans le dernier chapitre, nous utilisons la méthode faisceau pour résoudre un problème d'équilibre. Nous décrivons tout d'abord un algorithme général et ensuite un algorithme faisceau. Pour terminer, nous appliquons cette méthode aux problèmes d'inéquations variationnelles généralisés et multivoques.

Ce travail est essentiellement basé sur les articles [3] et [6] et le livre [4].

Chapitre 1

Problème d'équilibre et problèmes apparentés

Soit C un sous-ensemble convexe fermé non vide de $\mathbb{R}^n$ et soit $f : C \times C \to \mathbb{R}$ une bifonction d'équilibre continue telle que $f(x, \cdot)$ soit convexe sur C pour tout $x \in C$.

Définition 1.0.1 *Une bifonction $f : C \times C \to \mathbb{R}$ est appelée une bifonction d'équilibre si $f(x,x) = 0$ pour tout $x \in C$.*

Le problème d'équilibre consiste à

$$(EP) \quad \text{trouver } x^* \in C \text{ tel que } f(x^*, y) \geq 0 \text{ pour tout } y \in C.$$

Nous définissons le problème dual d'équilibre comme suit :

$$(DEP) \quad \text{trouver } y^* \in C \text{ tel que } f(x, y^*) \leq 0 \text{ pour tout } x \in C.$$

Le problème d'équilibre (EP) est très général dans le sens qu'il inclut, comme cas particuliers :

- le problème d'optimisation,
- le problème d'inéquation variationnelle,
- le problème d'équilibre de Nash dans les jeux non coopératifs,
- le problème de point fixe,
- le problème de complémentarité non linéaire,
- le problème d'optimisation vectoriel.

L'intérêt de ce problème est qu'il unifie tous ces problèmes particuliers de manière simple. De plus, beaucoup de méthodes destinées à résoudre un de ces problèmes peuvent être étendues, avec des modifications adaptées, pour résoudre le problème d'équilibre général.

La plupart des méthodes pour résoudre les problèmes d'équilibre sont dérivés des formulations du point fixe de (EP). Puisque $f(x,x) = 0$ pour tout $x \in C$, $x^* \in C$ est une solution du problème (EP) si et seulement si x^* est une solution du problème $\min_{x \in C} f(x^*, x)$. Alors, avec $x^0 \in C$ donné, l'algorithme correspondant génère une suite $\{x^k\}$ définie, pour tout $k \in \mathbb{N}$, par

$$x^{k+1} = \arg\min_{x \in C} f(x^k, x). \tag{1.1}$$

Cependant, il est plus commode d'utiliser, comme problèmes d'inéquations variationnelles, le principe du problème auxiliaire basé sur la propriété du point fixe :

$x^* \in C$ est une solution du problème (EP) si et seulement si x^* est une solution de

$$\min_{y \in C} \{\epsilon f(x^*, y) + h(y) - h(x^*) - \langle \nabla h(x^*), y \rangle\} \tag{1.2}$$

où $\epsilon > 0$ et $h : C \to \mathbb{R}$ est une fonction différentiable fortement convexe.

L'itération du point fixe correspondante est alors :

Soit $x^k \in C$ donné, trouver $x^{k+1} \in C$ la solution de

$$(P_k) \quad \min_{y \in C} \{\epsilon f(x^k, y) + h(y) - h(x^k) - \langle \nabla h(x^k), y \rangle\}.$$

Nous observons que ce problème a une solution unique puisque h est fortement convexe. Cet algorithme a été introduit par Mastroeni [5]. Nous prouverons sa convergence sous les hypothèses que f est fortement monotone sur $C \times C$, dans le sens qu'il existe $\gamma > 0$ tel que

$$f(x,y) + f(y,x) \leq -\gamma \|y - x\|^2 \quad \forall x, y \in C, \tag{1.3}$$

et que f satisfait la propriété : il existe $c, d > 0$ tels que

$$\forall x, y, z \in C \quad f(x,y) + f(y,z) \geq f(x,z) - c\|y - x\|^2 - d\|z - y\|^2. \tag{1.4}$$

Quand

$$f(x,y) = \langle F(x), y - x \rangle + \varphi(y) - \varphi(x) \quad \forall x, y \in C \tag{1.5}$$

avec $F : C \to \mathbb{R}^n$ un opérateur continu et $\varphi : C \to \mathbb{R}$ une fonction convexe continue, le problème (EP) est réduit au problème d'inéquation variationnelle généralisé

$(GVIP)$ Trouver $x^* \in C$ tel que, pour tout $y \in C$,

$$\langle F(x^*), y - x^* \rangle + \varphi(y) - \varphi(x) \geq 0. \tag{1.6}$$

Dans ce cas, l'algorithme du principe du problème auxiliaire devient :

Soit $x^k \in C$ donné, trouver $x^{k+1} \in C$ la solution du problème

$$\min_{y \in C} \epsilon[\varphi(y) + \langle F(x^k), y - x^k \rangle] + h(y) - h(x^k) - \langle \nabla h(x^k), y - x^k \rangle. \tag{1.7}$$

Il est possible de montrer que (1.3) et (1.4) sont satisfaits quand F est fortement monotone et Lipschitz continu, respectivement. Cependant ces hypothèses sont trop fortes dans ce cas. En effet, Zhu et Marcotte [8] ont prouvé que la suite $\{x^k\}$ générée par le principe du problème auxiliaire converge vers une solution du problème $(GVIP)$ quand F est co-coercif sur C dans le sens que

$$\exists \gamma > 0 \quad \forall x, y \in C \quad \langle F(y) - F(x), y - x \rangle \geq \gamma \|F(y) - F(x)\|^2. \tag{1.8}$$

Il est évident que F co-coercif sur C n'implique pas, en général, que la fonction correspondante f définie en (1.5) soit fortement monotone sur $C \times C$ (par exemple, si $F = 0$ nous observons que $f(x, y) + f(y, x) = 0$). Un des buts est donc d'obtenir la convergence de l'algorithme de Mastroeni sous les hypothèses plus faibles que (1.3) et (1.4) de telle manière que le résultat de Zhu et Marcotte puisse être dérivé comme un cas particulier.

En ce qui concerne l'implémentation de l'algorithme précédent, le sous-problème (P_k) peut être difficile à résoudre quand la fonction convexe $f(x^k, \cdot)$ est non linéaire. C'est le cas lorsque f est donnée par (1.5) et que φ est une fonction convexe non différentiable. Dans ce cas, notre stratégie est d'approximer la fonction $f(x^k, \cdot)$ par une autre fonction convexe donc que le sous-problème (P_k) devienne facile à résoudre et la convergence est préservée sous les mêmes hypothèses que dans le cas exact.

Cette stratégie a été utilisée par Konnov à un niveau inférieur d'une méthode de relaxation combinée (CRM) pour trouver des points d'équilibre. Plus précisément, soit $x^k \in C$ donné, Konnov considère des linéarisations successives de la fonction $f(x^k, \cdot)$ afin de construire une approximation linéaire par morceaux convexe $\overline{f}_k$ de $f(x^k, \cdot)$ telle que la solution y^k du sous-problème (P_k) avec $f(x^k, \cdot)$ remplacé par $\overline{f}_k$ satisfasse la propriété :

$$f(x^k, y^k) \leq \mu \overline{f}_k(y^k) \quad (0 < \mu < 1). \tag{1.9}$$

Nous considérons à présent le problème d'inéquation variationnelle multivoque :

$(MVIP)$ Trouver $x^* \in C$ et $r^* \in F(x^*)$ tels que, pour tout $y \in C$,

$$\langle r^*, y - x^* \rangle \geq 0$$

où C est un sous-ensemble convexe fermé non vide de $\mathbb{R}^n$ et $F : C \to 2^{\mathbb{R}^n}$ est un opérateur multivoque continu co-coercif à valeurs compactes.

Nous définissons le problème dual d'inéquation variationnelle multivoque comme suit :

$(DMVIP)$ trouver $x^* \in C$ tel que pour tout $y \in C$ et pour tout $r \in F(x)$,

$$\langle r, y - x^* \rangle \geq 0.$$

Le problème $(MVIP)$ est un exemple particulier du problème (EP) quand la fonction f est définie, pour tout $x, y \in C$, par

$$f(x,y) = \sup_{\xi \in F(x)} \langle \xi, y - x \rangle.$$

Dans le cas où $C = \mathbb{R}^n$, le problème $(MVIP)$ se réduit au problème d'inclusion multivoque suivant :

(MIP) Trouver $x^* \in \mathbb{R}^n$ tel que $0 \in F(x^*)$.

Nous définissons le problème dual d'inclusion multivoque comme suit :

$(DMIP)$ trouver $x^* \in \mathbb{R}^n$ tel que pour tout $x \in \mathbb{R}^n$ et pour tout $t \in F(x)$,

$$\langle t, x - x^* \rangle \geq 0.$$

Chapitre 2

Méthode de relaxation combinée pour le problème (EP)

2.1 Schéma de base pour construire la méthode de relaxation combinée

Soit C un sous-ensemble convexe non vide de l'espace $\mathbb{R}^n$. Le problème est de trouver un point x^* d'un certain sous ensemble $C^* \subset C$. Pour cet algorithme général, nous considérons un schéma de base pour construire des méthodes de relaxation combinée et certaines de ses propriétés. Dans ce but, nous introduisons les définitions et notations suivantes.

Définition 2.1.1 *Un opérateur $P : \mathbb{R}^n \to \mathbb{R}^n$ est dit*

- *admissibile par rapport à l'ensemble C si*

$$P(x) \in C$$

pour tout $x \in \mathbb{R}^n$.

- *quasi non expansif par rapport à l'ensemble C si, pour chaque point $x \in \mathbb{R}^n$, nous avons*

$$\|P(x) - z\| \leq \|x - z\| \quad \textit{pour tout } z \in C. \tag{2.1}$$

- *non expansif par rapport à l'ensemble C si, pour chaque paire de points $x, z \in C$, nous avons*

$$\|P(x) - P(z)\| \leq \|x - z\|.$$

Nous notons par $\mathcal{F}(C)$ l'ensemble des opérateurs admissibles quasi non expansifs par rapport à l'ensemble C.

Pour chaque point $x \in \mathbb{R}^n$ et pour chaque ensemble $V \subset \mathbb{R}^n$, nous posons

$$dist(x, V) = \inf_{y \in V} \|x - y\|.$$

Schéma de base 2.1.1 *Choisir un point $x^0 \in C$ et une suite $\{\gamma_j\}$ telle que*

$$\gamma_j \in [0, 2],\ j = 0, 1, \ldots;\quad \sum_{j=0}^{+\infty} \gamma_j (2 - \gamma_j) = +\infty. \tag{2.2}$$

Poser $\lambda(0) = 1$ et $j(0) = 0$.

A la k-ième itération, $k = 0, 1, \ldots$, nous avons un point $x^k \in C$ et les nombres $\lambda(k)$ et $j(k)$.

1. *Poser $l = \lambda(k)$, appliquer la procédure D_k avec le point d'entrée x^k et un nombre l et obtenir le vecteur de sortie g^k et les nombres ω_k et $\lambda(k+1)$.*
2. *Si $\lambda(k+1) > l$, poser $x^{k+1} = u^l = x^k$, $\sigma_k = 0$ et $j(k+1) = j(k)$. L'itération est terminée.*
3. *Si $\lambda(k+1) = l$,*

$$\sigma_k = \begin{cases} \omega_k / \|g^k\|^2 & \text{si } g^k \neq 0, \\ 0 & \text{sinon}\,; \end{cases}$$

$j(k+1) = j(k) + 1$, $x^{k+1} = P_k(x^k - \gamma_{j(k)} \sigma_k g^k)$, où $P_k \in \mathcal{F}(C)$. L'itération est terminée.

Selon cette description, le Schéma de base 2.1.1 contient la procédure D_k qui permet de calculer la direction de descente g^k, la longueur de pas σ_k ainsi que l'opérateur P_k. Les valeurs $\lambda(k)$ et $j(k)$ sont les compteurs de pas nuls et sérieux, respectivement.

Lemme 2.1.1 *Soit $\{x^k\}$ une suite construite suivant le Schéma de base 2.1.1 et supposons que la relation suivante ait lieu*

$$\langle g^k, x^k - x^* \rangle \geq \omega_k \geq 0 \quad \forall x^* \in C_k^* \subseteq C^*, \quad k = 0, 1, \ldots \tag{2.3}$$

Alors

1. $\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - \gamma_{j(k)}(2 - \gamma_{j(k)})(\sigma_k \|g^k\|)^2$

$$\forall x^* \in C_k^*, \quad k = 0, 1, \ldots \tag{2.4}$$

2. s'il existe la valeur d'adhérence x^* de $\{x^k\}$ qui appartient à l'ensemble $V_0 = \bigcap_{k=0}^{+\infty} C_k^*$, alors $\lim_{k \to +\infty} x^k = x^*$.

Preuve.

1. Soit x^* un point quelconque dans C_k^*.

   - Si $\lambda(k+1) > \lambda(k)$, alors $x^{k+1} = x^k$ et $\sigma_k = 0$. Par conséquent, (2.4) est vérifié.

   - Considérons maintenant le cas où $\lambda(k+1) = \lambda(k)$. Par définition de x^{k+1}, par le fait que P^k soit quasi non expansif, par (2.3) et par la définition de ω_k, nous avons alors

$$\begin{aligned}
\|x^{k+1} - x^*\|^2 &= \|P_k(x^k - \gamma_{j(k)}\sigma_k g^k) - x^*\|^2 \\
&\leq \|x^k - \gamma_{j(k)}\sigma_k g^k - x^*\|^2 \\
&= \|x^k - x^*\|^2 + (\gamma_{j(k)}\sigma_k \|g^k\|)^2 - 2\gamma_{j(k)}\sigma_k \langle g^k, x^k - x^* \rangle \\
&\leq \|x^k - x^*\|^2 + (\gamma_{j(k)}\sigma_k \|g^k\|)^2 - 2\gamma_{j(k)}\sigma_k \omega_k \\
&\leq \|x^k - x^*\|^2 - \gamma_{j(k)}(2 - \gamma_{j(k)})(\sigma_k \|g^k\|)^2
\end{aligned}$$

   i.e. (2.4) a lieu.

2. Suit de (2.4). □

Lemme 2.1.2 *Supposons que toutes les hypothèses du Lemme 2.1.1 soient satisfaites et supposons que soit $V_0 \neq \emptyset$, ou, pour chaque k, il existe un élément $x^* \in C_k^*$ tel que*

$$dist(x^k, C^*) = \|x^k - x^*\|, \quad k = 0, 1, \ldots$$

Alors

1. $\sum_{j=0}^{+\infty} \gamma_{j(k)}(2 - \gamma_{j(k)})(\sigma_k \|g^k\|)^2 < +\infty$ *;*

2. si

$$\omega_k \geq \overline{\omega} > 0 \quad \textit{lorsque} \quad \lambda(k) \leq \bar{l} < +\infty \tag{2.5}$$

et si $\{g^k\}$ est bornée, alors la suite $\{u^l\}$ est infinie.

Preuve.

1. Soit

$$a_k \equiv \gamma_{j(k)}(2 - \gamma_{j(k)})(\sigma_k \|g^k\|)^2 \geq 0.$$

Par (2.4), nous avons

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - a_k$$

et donc

$$0 \leq a_k \leq \|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2$$

où $k = 1, \ldots, n$.

En prenant la somme sur $k = 1$ jusqu'à n, nous obtenons

$$0 \leq \sum_{k=1}^{n} a_k \leq \|x^1 - x^*\|^2 - \|x^{n+1} - x^*\|^2$$

ou encore

$$0 \leq \sum_{k=1}^{n} a_k \leq \|x^1 - x^*\|^2.$$

En faisant tendre n vers l'infini, nous avons que $\sum_{k=1}^{+\infty} a_k$ est une série convergente.

2. – Montrons tout d'abord que $\liminf_{k\to+\infty}(\sigma_k \|g^k\|) = 0$

Procédons par l'absurde. Si

$$\sigma \equiv \lim_{n\to+\infty} \inf_{k\geq n} \sigma_k \|g^k\| > 0$$

alors il existe n_0 tel que $\forall n \geq n_0$

$$\inf_{k\geq n} \sigma_k \|g^k\| > \frac{\sigma}{2}.$$

En particulier (en prenant $n = n_0$), il existe n_0 tel que $\forall k \geq n_0$,

$$(\sigma_k \|g^k\|)^2 > \frac{\sigma^2}{4}.$$

Donc

$$\gamma_k(2-\gamma_k) = \gamma_k(2-\gamma_k)\frac{(\sigma_k\|g^k\|)^2}{(\sigma_k\|g^k\|)^2} < \frac{4}{\sigma^2}\gamma_k(2-\gamma_k)(\sigma_k\|g^k\|)^2.$$

Puisque

$$\sum_{j=0}^{+\infty} \gamma_{j(k)}(2-\gamma_{j(k)})(\sigma_k\|g^k\|)^2 < +\infty$$

nous déduisons que

$$\sum_{j=0}^{+\infty} \gamma_{j(k)}(2-\gamma_{j(k)}) < +\infty$$

qui contredit (2.2). Nous avons donc $\liminf_{k\to+\infty}(\sigma_k\|g^k\|) = 0$.

– Supposons par contradiction que sous les hypothèses de (2), la suite $\{u^l\}$ soit finie. Alors, en vertu du caractère borné de $\{\lambda(k)\}$ et $\{g^k\}$, nous avons

$$\begin{aligned}\liminf_{k\to+\infty} \omega_k &= \liminf_{k\to+\infty} \sigma_k\|g^k\|^2 \\ &= \underbrace{\liminf_{k\to+\infty} \sigma_k\|g^k\|}_{=0}\ \underbrace{\|g^k\|}_{\text{bornée}} \\ &= 0\end{aligned}$$

qui contredit (2.5). □

Il suit des propriétés obtenues que, en particulier, la procédure D_k a pour but que les relations (2.3) et (2.5) soient satisfaites, ainsi que l'existence d'une valeur d'adhérence de $\{x^k\}$ appartenant à $V_0 \subset C^*$.

2.2 CRM pour trouver des points d'équilibre sujets à des contraintes linéaires

Nous considérons le problème d'équilibre (EP) i.e. le problème qui consiste à trouver un point de l'ensemble

$$C^* = \{x^* \in C | f(x^*, y) \geq 0 \quad \forall y \in C\}$$

où $f : C \times C \to \mathbb{R}$ est une bifonction d'équilibre continue et $f(x, \cdot)$ est une fonction convexe pour chaque $x \in C$. Nous supposons également que $f(\cdot, y)$ est une fonction concave pour chaque $y \in C$.

Soit l'ensemble C défini aux moyens de fonctions linéaires, i.e.,

$$C = \{x \in \mathbb{R}^n | h_i(x) = \langle a_i, x\rangle - b_i \leq 0, \quad i \in M\},$$

$M = \{1, \ldots, m\}$, $a_i \in \mathbb{R}^n$, $b_i \in \mathbb{R}$, $i \in M$.

Dans ce qui suit, nous supposons que l'ensemble des solutions C^* est non vide. En prenant le Schéma de base 2.1.1 et les propriétés des méthodes de linéarisation comme base, nous décrivons une méthode itérative pour trouver un point de l'ensemble C^*.

Nous définissons à présent le sous-gradient et la sous-différentielle d'une fonction convexe.

Définition 2.2.1 *La sous-différentielle d'une fonction convexe $\phi : \mathbb{R}^n \to \mathbb{R}$ en $x \in \mathbb{R}^n$ est l'ensemble*

$$\partial\phi(x) = \{x^* \in \mathbb{R}^n | \phi(y) \geq \phi(x) + \langle x^*, y - x\rangle \text{ pour tout } y \in \mathbb{R}^n\}.$$

Les éléments de $\partial\phi(x)$ sont appelés sous-gradients de ϕ en x.

Soient x^k, y^k, et w^t des points de C. Nous introduisons les notations suivantes

$$\begin{aligned}
f_k(y) &= f(x^k, y);\\
\psi_k(x) &= f(x, y^k);\\
\varphi_{k,t}(y) &= f_k(w^t) + \langle q^t, y - w^t\rangle, \quad q^t \in \partial f_k(w^t);\\
\overline{\varphi}_{k,t} &= \varphi_{k,t}(x^k), \quad i \in M;\\
\overline{h}_i &= h_i(x^k), \quad i \in M.
\end{aligned}$$

Méthode 2.2.1 *Choisir un nombre $\mu \in (0,1)$, une suite matricielle $\{A_k\}$ telle que*

$$\lambda_{min}\|p\|^2 \leq \langle A_k p, p\rangle \leq \lambda_{max}\|p\|^2 \quad \forall p \in I\!R^n, \quad 0 < \lambda_{min} \leq \lambda_{max} < +\infty \tag{2.6}$$

et une suite $\{\epsilon_l\} \searrow 0$.

La méthode consiste à appliquer le Schéma de base 2.1.1 où la procédure D_k est décrite comme suit.

Données *: un point x^k et un nombre l.*

Sorties *: un élément g^k et des nombres ω_k et $\lambda(k+1)$.*

Pas 1. *Poser $s = 0$, $w^0 = x^k$, et $J_0 = \{0\}$, choisir $q^0 \in \partial f_k(\omega^0)$ et poser $\overline{\varphi}_{k,0} = f_k(\omega^0)$.*

Pas 2. *Trouver le vecteur $d^s \in I\!R^n$ et le nombre β_s qui résolvent le problème d'optimisation suivant :*

$$\min \quad \beta + \frac{1}{2}\langle A_k d, d\rangle \tag{2.7}$$

$$\text{s.c.} \quad \overline{\varphi}_{k,t} + \langle q^t, d\rangle \leq \beta, \quad t \in J_s, \tag{2.8}$$

$$\overline{h}_i + \langle a^i, d\rangle \leq 0, \quad i \in M. \tag{2.9}$$

Pas 3. *Si $\beta_s \geq -\epsilon_l$, alors poser $g^k = 0$, $\omega_k = 0$ et $\lambda(k+1) = l+1$ et STOP.*

Pas 4. *Poser $w^{s+1} = w^0 + d^s$. Si*

$$f_k(w^{s+1}) \leq \mu\beta_s, \tag{2.10}$$

alors poser $y^k = w^{s+1}$, $\omega_k = -f_k(y^k)$ et $\lambda(k+1) = l$, choisir $g^k \in \partial(-\psi_k(x^k))$ et STOP.

Pas 5. *Si $f_k(w^{s+1}) > \mu\beta_s$, poser $J_{s+1} = J_s \bigcup \{s+1\}$, calculer $q^{s+1} \in \partial f_k(w^{s+1})$,*

$$\overline{\varphi}_{k,s+1} = f_k(w^{s+1}) + \langle q^{s+1}, w^0 - w^{s+1}\rangle$$

et retour au Pas 2.

Définition 2.2.2 *Nous appelons un pas intérieur un changement de l'indice s dans la procédure D_k.*

Lemme 2.2.1 *La solution (β_s, d^s) du problème (2.7)-(2.9) existe toujours et est unique, et nous avons*

$$\beta_s + \frac{1}{2}\langle A_k d^s, d^s\rangle \leq 0. \tag{2.11}$$

Preuve.

– Le point $(0,0) \in \mathrm{I\!R} \times \mathrm{I\!R}^n$ satisfait les contraintes (2.8) et (2.9)

Par définitions et car $f_k(\cdot)$ est convexe, nous avons en effet

$$\begin{aligned} \overline{\varphi}_{k,t} &= \varphi_{k,t}(x^k) \\ &= f_k(w^t) + \langle q^t, x^k - w^t\rangle \\ &\leq f_k(x^k) \\ &= f(x^k, x^k) \\ &= 0, \quad t \in J_s. \end{aligned}$$

Par définition et car $x^k \in C$, nous avons, pour $i \in M$,

$$\begin{aligned} \overline{h}_i &= h_i(x^k) \\ &= \langle a_i, x^k\rangle - b_i \leq 0. \end{aligned}$$

– Unicité

En vertu de (2.6), la fonction coût $\beta + \frac{1}{2}\langle A_k d, d\rangle$ est fortement convexe en d. La valeur optimale β_s étant uniquement déterminée par d^s en vertu de (2.8), il suit alors que la solution (β_s, d^s) est unique.

– Existence

Suit du fait que, pour tous β et d admissibles, les ensembles

$$\left\{(\beta, d) | \beta + \frac{1}{2}\langle A_k d, d\rangle \leq \overline{\beta}\right\}$$

soient bornés.

– La relation (2.11) suit du fait que

$$\beta_s + \frac{1}{2}\langle A_k d_s, d_s\rangle \leq \beta + \frac{1}{2}\lambda_{max}\|d\|^2$$

et que le point $(0,0)$ est admissible. □

Lemme 2.2.2 *1. Le nombre de pas intérieurs dans la procédure D_k est fini.*

2. Si la méthode se termine avec $\lambda(k+1) = l+1$, alors

$$f_k(y) \geq -\epsilon_l - \|y - \omega_0\| \sqrt{\frac{2\epsilon_l}{\tau'}} \quad \forall y \in C, \tag{2.12}$$

où

$$\tau' \leq \frac{\lambda_{min}}{(\lambda_{max})^2}. \tag{2.13}$$

Preuve.

1. Supposons, par contradiction, que les suites $\{\beta_s\}$ et $\{d^s\}$ soient infinies. Par le Lemme 2.2.1, nous avons

$$\beta_s + \frac{1}{2}\langle A_k d^s, d^s\rangle \leq 0$$

et par (2.8),

$$\overline{\varphi}_{k,t} + \langle q^t, d^s\rangle \leq \beta_s \quad \text{pour tout } t \in J_s.$$

Soit $t \in J_s$. En $t = 0$, nous avons

$$\overline{\varphi}_{k,0} + \langle q^0, d^s\rangle + \frac{1}{2}\|d^s\|_A^2 \leq 0.$$

Par Cauchy-Schwarz et (2.6),

$$\overline{\varphi}_{k,0} - \|q^0\|\,\|d^s\| + \frac{1}{2}\lambda_{\min}\|d^s\|^2 \leq 0$$

ou encore

$$\overline{\varphi}_{k,0} + \|d^s\| \left[-\|q^0\| + \frac{1}{2}\lambda_{\min}\|d^s\| \right] \leq 0$$

ce qui n'est pas possible.

Nous avons donc $\|d^s\| \leq C_1 < +\infty$ et $\beta_s \to -\infty$.

Par (2.8) en $t = 0$, nous avons

$$\begin{aligned} \beta_s &\geq \overline{\varphi}_{k,0} - \|q^0\|\,\|d^s\| \\ &\geq \underbrace{\overline{\varphi}_{k,0} - C_1\|q^0\|}_{\in \mathbb{R}} \end{aligned}$$

ce qui n'est pas possible.

Par conséquent, sans perdre de généralité, nous concluons que

$$\lim_{s\to+\infty} \beta_s = \overline{\beta} \leq -\epsilon_l < 0, \quad \lim_{s\to+\infty} d^s = \overline{d}. \tag{2.14}$$

Pour chaque s, nous avons

$$\begin{aligned} f_k(w^{s+1}) &\equiv f_k(w^{s+1}) + \langle q^{s+1}, w^0 - w^{s+1}\rangle + \langle q^{s+1}, w^{s+1} - w^0\rangle \\ &= \overline{\varphi}_{k,s+1} + \langle q^{s+1}, d^s\rangle. \end{aligned}$$

Mais si la méthode ne se termine pas, alors (2.10) n'a pas lieu, et par conséquent

$$\overline{\varphi}_{k,s+1} + \langle q^{s+1}, d^s\rangle > \mu\beta_s$$

et, par (2.8), nous avons

$$\overline{\varphi}_{k,s+1} + \langle q^{s+1}, d^{s+1}\rangle \leq \beta_{s+1}.$$

Par conséquent,

$$\langle q^{s+1}, d^{s+1} - d^s\rangle \leq \beta_{s+1} - \mu\beta_s.$$

Puisque la suite $\{d^s\}$ est bornée, $\{q^s\}$ l'est aussi. Par conséquent, en passant à la limite lorsque $s \to +\infty$ dans la relation ci-dessus, nous obtenons

$$\begin{aligned} 0 &= \lim_{s\to+\infty} \langle q^{s+1}, d^{s+1} - d^s\rangle \\ &\leq \lim_{s\to+\infty} (\beta_{s+1} - \mu\beta_s) \\ &= (1-\mu)\overline{\beta} \end{aligned}$$

qui contredit (2.14). Par conséquent, l'assertion (1) est valide.

2. La fonction Lagrangienne associée au problème (2.7)-(2.9) est

$$L(\beta, d, \lambda, \mu) = \beta + \frac{1}{2}\langle A_k d, d\rangle + \sum_{t\in J_s} \lambda_t[\overline{\varphi}_{k,t} + \langle q^t, d\rangle - \beta] + \sum_{i\in M} \mu_i[\overline{h}_i + \langle a_i, d\rangle].$$

La fonction duale est $d(\lambda,\mu) = \min_{\beta,d} L(\beta, d, \lambda, \mu)$. Pour trouver le minimum de $L(\beta, d, \lambda, \mu)$, nous résolvons

$$\begin{aligned} \nabla L_\beta(\beta, d, \lambda, \mu) = 0 &\Leftrightarrow 1 - \sum_{t\in J_s} \lambda_t = 0 \\ &\Leftrightarrow \sum_{t\in J_s} \lambda_t = 1 \end{aligned}$$

et

$$\begin{aligned} \nabla L_d(\beta, d, \lambda, \mu) = 0 &\Leftrightarrow A_k d + \sum_{t\in J_s} \lambda_t q^t + \sum_{i\in M} \mu_i a_i = 0 \\ &\Leftrightarrow \sum_{t\in J_s} \lambda_t q^t + \sum_{iM} \mu_i a_i = \underbrace{-A_k d}_{\equiv\widetilde{d}}. \end{aligned}$$

La fonction duale devient

$$\begin{aligned}
d(\lambda,\mu) &= \beta + \frac{1}{2}\langle A_k d, d\rangle + \sum_{t\in J_s}\lambda_t\overline{\varphi}_{k,t} + \sum_{i\in M}\mu_i\overline{h}_i \\
&\quad -\beta + \left\langle \sum_{t\in J_s}\lambda_t q^t + \sum_{i\in M}\mu_i a^i, d\right\rangle \\
&= \frac{1}{2}\langle \widetilde{d}, A_k^{-1}\widetilde{d}\rangle + \sum_{t\in J_s}\lambda_t\overline{\varphi}_{k,t} + \sum_{i\in M}\mu_i\overline{h}_i - \langle \widetilde{d}, A_k^{-1}\widetilde{d}\rangle \\
&= \sum_{t\in J_s}\lambda_t\overline{\varphi}_{k,t} + \sum_{i\in M}\mu_i\overline{h}_i - \frac{1}{2}\langle \widetilde{d}, A_k^{-1}\widetilde{d}\rangle.
\end{aligned}$$

Le problème dual associé à (2.7)-(2.9) est donc :

$$\max \quad \sum_{t\in J_s}\lambda_t\overline{\varphi}_{k,t} + \sum_{i\in M}\mu_i\overline{h}_i - \frac{1}{2}\langle \widetilde{d}, A_k^{-1}\widetilde{d}\rangle \tag{2.15}$$

$$s.c. \quad \widetilde{d} = \sum_{t\in J_s}\lambda_t q^t + \sum_{i\in M}\mu_i a_i, \tag{2.16}$$

$$\sum_{t\in J_s}\lambda_t = 1, \quad \lambda_t \geq 0, \quad t\in J_s, \tag{2.17}$$

$$\mu_i \geq 0, \quad i\in M. \tag{2.18}$$

Soit (λ^*, μ^*) la solution du problème dual.

Choisissons un point quelconque $y \in C$ et posons $d = y - w^0$. Alors, pour chaque $t \in J_s$,

$$\begin{aligned}
f_k(y) &\geq f_k(w^t) + \langle q^t, y - w^t\rangle \\
&= f_k(w^t) + \langle q^t, w^0 - w^t\rangle + \langle q^t, y - w^0\rangle \\
&= \varphi_{k,t}(w^0) + \langle q^t, d\rangle \\
&= \overline{\varphi}_{k,t} + \langle q^t, d\rangle
\end{aligned}$$

et pour tout $i \in M$,

$$\begin{aligned}
0 &\geq h_i(w^0 + d) \\
&= \langle a_i, w^0 + d\rangle - b_i \\
&= \langle a_i, w^0\rangle - b_i + \langle a_i, d\rangle \\
&= \overline{h}_i + \langle a_i, d\rangle.
\end{aligned}$$

En multipliant la t-ième inégalité par λ_t^* et la i-ième inégalité par μ_i^* et en additionnant toutes ces inégalités nous obtenons

$$f_k(y) \geq \sum_{t\in J_s}\lambda_t^*\overline{\varphi}_{k,t} + \sum_{i\in M}\mu_i^*\overline{h}_i + \left\langle \sum_{t\in J_s}\lambda_t^* q^t + \sum_{i\in M}\mu_i^* a_i, d\right\rangle.$$

Puisque les valeurs optimales du problème primal et du problème dual coïncident, nous avons

$$f_k(y) \geq \beta_s + \frac{1}{2}\langle A_k d^s, d^s\rangle + \frac{1}{2}\langle \widetilde{d^s}, A_k^{-1}\widetilde{d^s}\rangle + \langle \widetilde{d^s}, d\rangle$$

où $\widetilde{d^s} = \sum_{t\in J_s} \lambda_t^* q^t + \sum_{i\in M} \mu_i^* a_i = -A_k d^s$.

D'où $f_k(y) \geq \beta_s - \|d\| \, \|\widetilde{d^s}\|$.

En prenant en compte (2.6), (2.13) et (2.11), nous avons

$$\|\widetilde{d^s}\|^2 = \|A_k d^s\|^2 \leq \lambda_{max}^2 \|d^s\|^2 \leq \frac{\lambda_{max}^2}{\lambda_{min}}\langle A_k d^s, d^s\rangle \leq \frac{\langle A_k d^s, d^s\rangle}{\tau'} \leq -\frac{2\beta_s}{\tau'}.$$

Mais la relation $\lambda(k+1) = l+1$ implique $\beta_s \geq -\epsilon_l$ pour un certain s, par conséquent

$$\|\widetilde{d^s}\|^2 \leq 2\epsilon_l/\tau'$$

et

$$f_k(y) \geq -\epsilon_l - \|y - \omega_0\| \sqrt{\frac{2\epsilon_l}{\tau'}}.$$

□

Lemme 2.2.3 *Dans la procédure D_k, la relation $\lambda(k+1) = l$ implique $g^k \neq 0$.*

Preuve.
Supposons le contraire, i.e. $\lambda(k+1) = l$ et $g^k = 0$.

Par définition de g^k, il suit que

$$-\psi_k(x) \geq -\psi_k(x^k) + \langle g^k, x - x^k\rangle.$$

Par définition de $\psi_k(\cdot)$ et car $g^k = 0$, la relation précédente est équivalente à

$$-f(x, y^k) \geq -f(x^k, y^k)$$

et donc

$$f(x^k, y^k) \geq f(x, y^k)$$

pour chaque $x \in C$. Par conséquent, dans le cas où $x = y^k$, nous avons $f(x^k, y^k) \geq 0$ car f est une bifonction d'équilibre et donc $f(y^k, y^k) = 0$ pour tout $y \in C$.

D'un autre côté, par (2.10) nous avons

$$f(x^k, y^k) = f_k(w^{s+1}) \leq \mu\beta_s < -\mu\epsilon_l < 0.$$

Nous obtenons donc une contradiction. □

Du Lemme 2.2.3, il suit que dans le Schéma de base 2.1.1 au Pas 3, nous avons toujours $g^k \neq 0$ et donc $\sigma_k = \omega_k / \|g^k\|^2 > 0$.

2.3 Convergence

Nous prouvons à présent la convergence de la Méthode 2.2.1 en utilisant les propriétés obtenues précédemment.

Théorème 2.3.1 *Soit $\{x^k\}$ une suite construite par la Méthode 2.2.1. Alors*

1. *le nombre de pas intérieurs à chaque itération de la méthode est fini ;*
2. $\lim_{k\to+\infty} x^k = x^* \in C^*$.

Preuve.

1. Suit du Lemme 2.2.2 (1).

2. – Montrons qu'à la k-ième itération de la méthode nous avons

$$\langle g^k, x^k - x^* \rangle \geq \omega_k \geq 0 \quad \forall x^* \in C^* \tag{2.19}$$

Si $\lambda(k+1) > \lambda(k)$, alors $g^k = 0$, $w_k = 0$ et (2.19) a lieu.

Si $\lambda(k+1) = \lambda(k)$, alors $g^k \in \partial(-\psi_k(x^k))$ et $\omega_k = -f_k(y^k) \geq -\mu\beta_s > \mu\epsilon_l > 0$. Par conséquent, pour chaque $x^* \in C^*$, nous avons

$$-f(x^*, y^k) + f(x^k, y^k) \geq \langle g^k, x^* - x^k \rangle$$

d'où

$$\langle g^k, x^k - x^* \rangle \geq f(x^*, y^k) - f(x^k, y^k) \geq \omega_k$$

i.e., (2.19) a lieu.

– (2.5) est vérifié

Si, pour tout k, $\lambda(k) \leq \bar{l} < +\infty$, alors il existe $\bar{k}$ tel que pour tout $k \geq \bar{k}$ nous avons $\lambda(k+1) = \lambda(k)$ et, par construction de la procédure D_k,

$$\omega_k = -f(x^k, y^k) \geq -\mu\beta_s \geq \mu\epsilon_l > \mu\epsilon_{\bar{l}} \equiv \overline{\omega} > 0$$

i.e., (2.5) a lieu.

Toutes les hypothèses des Lemmes 2.1.1 et 2.1.2 sont donc satisfaites avec $V_0 = C^*$.

De (2.4), où x^* est un point arbitraire de C^*, il suit que la suite $\{x^k\}$ est bornée et donc $\{g^k\}$ l'est aussi car $g^k \in \partial(-\psi_k(x^k))$. Du Lemme 2.1.2 (2), il suit alors que la suite $\{u^l\}$ est infinie. Par conséquent, elle a une valeur d'adhérence. Sans perdre de généralité, nous pouvons conclure que $\lim_{l \to +\infty} u^l = x^* \in U$ parce que $u^l \in C$ et C est borné. Du Lemme 2.2.2 (2), il suit maintenant que, pour chaque $y \in C$,

$$f(u^l, y) \geq -\epsilon_l - \|y - u^l\| \sqrt{\frac{2\epsilon_l}{\tau'}}$$

où τ' est défini par (2.13). En passant à la limite lorsque $l \to +\infty$ et en prenant en compte la continuité de f, nous obtenons $f(x^*, y) \geq 0$, i.e., $x^* \in C^*$ en vertu du caractère arbitraire de $y \in C$. La deuxième partie du théorème suit alors du Lemme 2.1.1 (2). $\square$

2.4 Méthode combinée de non-descente pour trouver des points d'équilibre

D'autres schémas que le 2.1.1 peuvent être utilisés pour construire des méthodes combinées. Une telle approche autorise d'élargir les ressources dans le contrôle des paramètres de ces méthodes et d'accélérer la convergence du processus de calcul à cause de l'utilisation de propriétés spécifiques du problème sous solution. Dans cette section, nous considérons un des schémas de non-descente possible, qui est basé sur une généralisation de la méthode du sous-gradient. Nous décrivons tout d'abord ce schéma de base appliqué au problème pour trouver un point de l'ensemble C^* qui est contenu dans l'ensemble convexe $C \subset \mathbb{R}^n$.

Schéma de base 2.4.1 *Choisir un point $x^0 \in C$ et une suite $\{\alpha_j\}$ telle que*

$$\alpha_j > 0,\ j = 0, 1, \ldots; \quad \sum_{j=0}^{+\infty} \alpha_j = +\infty, \ \sum_{j=0}^{+\infty} \alpha_j^2 < +\infty. \tag{2.20}$$

Poser $\lambda(0) = 1$ et $j(0) = 0$.

A la k-ième itération, $k = 0, 1, \ldots$, nous avons un point $x^k \in C$ et les nombres $\lambda(k)$ et $j(k)$.

1. *Poser $l = \lambda(k)$, appliquer la procédure D_k avec le point d'entrée x^k et un nombre l et obtenir le vecteur de sortie g^k et les nombres ω_k et $\lambda(k+1)$.*
2. *Si $\lambda(k+1) > l$, poser $x^{k+1} = u^l = x^k$, $\sigma_k = 0$ et $j(k+1) = j(k)$. L'itération est terminée.*
3. *Si $\lambda(k+1) = l$,*

$$\sigma_k = \begin{cases} \alpha_{j(k)}/\|g^k\| & \text{si } g^k \neq 0, \\ 0 & \text{sinon;} \end{cases}$$

$j(k+1) = j(k) + 1$, $x^{k+1} = P_k(x^k - \sigma_k g^k)$, où $P_k \in \mathcal{F}(C)$. L'itération est terminée.

Lemme 2.4.1 *Soit $\{x^k\}$ une suite construite suivant le Schéma de base 2.4.1. Supposons que, pour $k = 0, 1, \ldots$, la relation (2.3) ait lieu, où l'ensemble $V_0 = \bigcap_{k=0}^{+\infty} C_k^*$ est non vide. Alors*

1. *la suite $\{x^k\}$ est bornée ;*
2. *si (2.5) a lieu et si la suite $\{g^k\}$ est bornée, alors la suite $\{u^l\}$ est infinie ;*
3. *s'il existe la valeur d'adhérence x^* de $\{x^k\}$ qui appartient à V_0, alors $\lim_{k\to+\infty} x^k = x^*$.*

Preuve.

1. Soit x^* un point quelconque dans V_0. En utilisant les définitions de x^{k+1} et P_k, (2.3) et la définition de ω_k, nous avons

$$\begin{aligned}
\|x^{k+1} - x^*\|^2 &= \|P_k(x^k - \sigma_k g^k) - x^*\|^2 \\
&\leq \|x^k - \sigma_k g^k - x^*\|^2 \\
&= \|x^k - x^*\|^2 - 2\sigma_k \langle g^k, x^k - x^* \rangle + (\sigma_k \|g^k\|)^2 \\
&\leq \|x^k - x^*\|^2 - 2\sigma_k \omega_k + (\sigma_k \|g^k\|)^2 \\
&\leq \|x^k - x^*\|^2 + (\sigma_k \|g^k\|)^2.
\end{aligned}$$

Par (2.20) nous avons aussi

$$\sum_{k=0}^{+\infty} (\sigma_k \|g^k\|)^2 \leq \sum_{j=0}^{+\infty} \alpha_j^2 < +\infty$$

et par conséquent (1) a lieu.

2. Supposons par contradiction que sous les hypothèses de (2), la suite $\{u^l\}$ soit finie. Nous avons alors

$$\begin{aligned}
\|x^{k+1} - x^*\|^2 &\leq \|x^k - x^*\|^2 - \sigma_k(2\omega_k - \sigma_k \|g^k\|^2) \\
&\leq \|x^k - x^*\|^2 - \sigma_k(2\overline{\omega} - \sigma_k \|g^k\|^2) \\
&= \|x^k - x^*\|^2 - \frac{\alpha_{j(k)}}{\|g^k\|}(2\overline{\omega} - \alpha_{j(k)} \|g^k\|) \\
&= \|x^k - x^*\|^2 - \frac{2\overline{\omega}\alpha_{j(k)}}{\|g^k\|} + \alpha_{j(k)}^2.
\end{aligned}$$

La suite $\{g^k\}$ étant bornée par hypothèse, nous avons

$$\|g^k\| \leq \frac{\overline{\omega}}{\alpha_{j(k)}} \quad \text{ou encore} \quad \alpha_{j(k)} \leq \frac{\overline{\omega}}{\|g^k\|}.$$

Nous obtenons alors

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - \overline{\omega}\alpha_{j(k)}/\|g^k\|$$

pour k suffisamment large.

Avec $M = \frac{\overline{w}}{\alpha_{j(k)}}$, nous avons

$$M\alpha_{j(k)} \leq \|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2$$

et en prenant la somme sur $k = \overline{k}$ jusqu'à l'infini

$$M \sum_{k=\bar{k}}^{+\infty} \alpha_{j(k)} \leq \|x^{\bar{k}} - x^*\|^2$$

ce qui contredit (2.20).

3. Du Lemme 1 de [1] (p. 160), il suit que $\lim_{k \to +\infty} \|x^k - x^*\| = \delta \geq 0$. Dans le cas où x^* est une valeur d'adhérence de $\{x^k\}$, nous devons avoir $\delta = 0$, d'où (3) a lieu. □

La Méthode 1.4.1 pour résoudre le problème d'équilibre défini au premier chapitre coïncide avec la Méthode 2.2.1, où le Schéma de base 2.1.1 est remplacé par le Schéma de base 2.4.1.

Théorème 2.4.1 *Soit $\{x^k\}$ une suite construite par la Méthode 1.4.1. Alors*

1. le nombre de pas intérieurs à chaque itération de la méthode est fini;

2. nous avons $\lim_{k \to +\infty} x^k = x^* \in C^*$.

Preuve.

1. Suit du Lemme 2.2.2 (1).
2. Les relations (2.19) et (2.5) sont prouvées par les mêmes arguments que ceux de la preuve du Théorème 2.3.1. Par le Lemme 2.4.1 (1), le caractère borné de $\{g^k\}$ peut être prouvé de manière analogue. Du Lemme 2.4.1 (2) il suit maintenant que la suite u^l a une valeur d'adhérence $x^* \in C$, d'où, sans perdre de généralité, $\lim_{l \to +\infty} u^l = x^*$. Aussi, la relation $x^* \in C^* = V_0$ est prouvée par les mêmes arguments que ceux de la preuve du Théorème 2.3.1. Par conséquent, par le Lemme 2.4.1 (3), nous concluons que la partie (2) du théorème a également lieu. □

Chapitre 3

Méthode de relaxation combinée pour le problème $(GVIP)$

Dans ce chapitre, nous décrivons une méthode de relaxation combinée, pour le problème $(GVIP)$. Tout au long de ce chapitre nous supposons que

- $F : V \to \mathbb{R}^n$ est un opérateur monotone et localement Lipschitz continu,
- $\varphi : \mathbb{R}^n \to \mathbb{R}$ est une fonction convexe,
- V est un ensemble fermé convexe dans $\mathbb{R}^n$ tel que $C \subseteq V$.

Nous supposons également que $(GVIP)$ est résolvable.

3.1 Description de la méthode

Nous définissons tout d'abord une fonction auxiliaire $h : V \to \mathbb{R}$ satisfaisant les hypothèses suivantes :

(H1) h est différentiable et son gradient ∇h est Lipschitz continu de constante $\tau'' < +\infty$;
(H2) h est fortement convexe de constante $\tau' > 0$.

Il est clair que la classe des fonctions satisfaisant (H1) et (H2) est plutôt large ; elle contient, en particulier, les fonctions quadratiques fortement convexes. Nous avons l'intention d'utiliser de telles fonctions pour construire une méthode CR pour résoudre $(GVIP)$ (3.3).

Méthode 3.1.1 *Choisir un point $x^0 \in V$, une suite $\{\gamma_k\}$ telle que*

$$\gamma_k \in [0,2],\ k = 0,1,\ldots; \quad \sum_{k=0}^{+\infty} \gamma_k(2-\gamma_k) = +\infty\,; \tag{3.1}$$

et une fonction $h : V \to I\!R$ satisfaisant (H1) et (H2). Choisir une suite d'opérateurs $\{P_k\}$, où $P_k \in \mathcal{F}(V_k)$ et $C \subseteq V_k \subseteq V$ pour $k = 0,1,\ldots$. Choisir des nombres $\alpha \in (0,1)$, $\beta \in (0,1)$ et $\widetilde{\theta} > 0$. Poser $k = 0$.

Pas 1 (Procédure auxiliaire $\overline{D}_k$)

Pas 1.1 : Trouver m le plus petit nombre dans $Z\!\!\!Z_+$ tel que

$$\langle F(x^k) - F(z^{k,m}), x^k - z^{k,m}\rangle \le (1-\alpha)(\widetilde{\theta}\beta^m)^{-1}\langle \nabla h(z^{k,m}) - \nabla h(x^k), z^{k,m} - x^k\rangle \tag{3.2}$$

où $z^{k,m}$ est une solution du problème d'optimisation auxiliaire suivant :

$$\min \quad \{\langle F(x^k) - (\widetilde{\theta}\beta^m)^{-1}\nabla h(x^k), y\rangle + (\widetilde{\theta}\beta^m)^{-1}h(y) + \varphi(y) | y \in C\}. \tag{3.3}$$

Pas 1.2 : Poser $\theta_k = \widetilde{\theta}\beta^m$ et $y^k = z^{k,m}$. Si $x^k = y^k$, STOP.

Pas 2 (Itération principale)

Poser

$$\begin{aligned} g^k &= F(y^k) - F(x^k) - \theta_k^{-1}(\nabla h(y^k) - \nabla h(x^k))\,; &(3.4)\\ \sigma_k &= \langle g^k, x^k - y^k\rangle / \|g^k\|^2\,; &(3.5)\\ x^{k+1} &= P_k(x^k - \gamma_k\sigma_k g^k)\,; &(3.6)\end{aligned}$$

$k = k+1$ et retour au Pas 1.

Nous notons $N(C,x)$ le cône normal à C en x défini par

$$N(C,x) = \{q \in \mathbb{R}^n | \langle q, y-x\rangle \le 0 \quad \forall y \in C\}$$

et $Nr(C)$ l'élément de C le plus proche de l'origine.

Nous considérons à présent les propriétés du problème (3.3) plus en détails.

Lemme 3.1.1 *1. Le problème (3.3) est équivalent à chacun des problèmes suivants :*

a. trouver $\overline{z} \in C$ *tel que*

$$\langle F(x^k), y - \overline{z}\rangle + (\widetilde{\theta}\beta^m)^{-1}[h(y) - h(\overline{z}) - \langle \nabla h(x^k), y - \overline{z}\rangle] + \varphi(y) - \varphi(\overline{z}) \geq 0 \quad \forall y \in C\,; \tag{3.7}$$

b. trouver $\overline{z} \in C$ *tel que*

$$\langle F(x^k), y - \overline{z}\rangle + (\widetilde{\theta}\beta^m)^{-1}\langle \nabla h(\overline{z}) - \nabla h(x^k), y - \overline{z}\rangle + \varphi(y) - \varphi(\overline{z}) \geq 0 \quad \forall y \in C\,; \tag{3.8}$$

c. trouver $\overline{z} \in C$ *tel que*

$$\exists \overline{g} \in \partial\varphi(\overline{z}) \,:\, \langle F(x^k) + \overline{g}, y - \overline{z}\rangle + (\widetilde{\theta}\beta^m)^{-1}\langle \nabla h(\overline{z}) - \nabla h(x^k), y - \overline{z}\rangle \geq 0 \quad \forall y \in C. \tag{3.9}$$

2. Le problème (3.3) a une solution unique.

Preuve.

1. <u>(3.3) $\Leftrightarrow$ (3.9).</u> Soit $\overline{z} \in C$ la solution du problème (3.3). Nous avons alors, par la Proposition A.0.1,

$$0 \in F(x^k) - (\widetilde{\theta}\beta^m)^{-1}\nabla h(x^k) + (\widetilde{\theta}\beta^m)^{-1}\nabla h(\overline{z}) + \partial\varphi(\overline{z}) + N(C, \overline{z})$$

ou encore

$$-[F(x^k) - (\widetilde{\theta}\beta^m)^{-1}\nabla h(x^k) + (\widetilde{\theta}\beta^m)^{-1}\nabla h(\overline{z}) + \partial\varphi(\overline{z})] \in N(C, \overline{z})$$

i.e. il existe $\overline{g} \in \partial\varphi(\overline{z})$ tel que

$$\langle F(x^k) + \overline{g} + (\widetilde{\theta}\beta^m)^{-1}\langle \nabla h(\overline{z}) - \nabla h(x^k), y - \overline{z}\rangle \geq 0 \quad \forall y \in C.$$

<u>(3.9) $\Leftrightarrow$ (3.8).</u> Suit de la Proposition B.0.1 en utilisant

$$F(x^k) + (\theta\beta^m)^{-1}(\nabla h(z) - \nabla h(x^k))$$

au lieu de $F(\overline{z})$.

<u>(3.9) $\Leftrightarrow$ (3.8).</u> Poser

$$\begin{aligned} f(x,y) &= \langle F(x^k), y - x\rangle + \varphi(y) - \varphi(x) \\ &\quad + (\theta\beta^m)^{-1}[h(y) - h(x) - \langle \nabla h(x^k), y - x\rangle]. \end{aligned}$$

Il est clair que f est une bifonction d'équilibre, que $f(x, \cdot)$ est convexe et sous-différentiable, et donc que

$$\partial f_y(x,y)|_{y=x} = F(x^k) + \partial\varphi(x) + (\theta\beta^m)^{-1}(\nabla h(x) - \nabla h(x^k)).$$

En appliquant le Théorème B.0.1 (2), nous pouvons conclure que (3.9) est équivalent à (3.7).

2. Suit du Corollaire A.0.1 puisque la fonction coût en (3.3) est fortement convexe.

$\square$

Nous pouvons donc remplacer le problème d'optimisation auxiliaire (3.3) dans la Méthode 3.1.1 par n'importe quel problème (3.7)-(3.9), qui sera plus pratique pour l'implémentation numérique. Néanmoins, le problème d'optimisation convexe (3.3) semble plus simple que $(MVIP)$ (3.9) ou $(GVIP)$ (3.7) et (3.8).

3.2 Convergence

Nous montrons tout d'abord que la procédure de recherche linéaire au Pas 1 est bien définie.

Lemme 3.2.1 *Nous avons*

1. $\theta_k > 0$, *i.e., la procédure de recherche linéaire dans le Pas 1 est toujours finie ;*
2. *si* $\{x^k\}$ *est bornée, alors* $\{y^k\}$ *l'est aussi et*

$$\theta_k \geq \theta' > 0, \quad k = 0, 1, \ldots \tag{3.10}$$

Preuve.

1. Soit $\widetilde{x} \in C$. En prenant en compte le Lemme 3.1.1 (1) et en appliquant (3.8) avec $y = \widetilde{x}$, $\overline{z} = z^{k,m}$ nous obtenons

$$\langle F(x^k) + (\widetilde{\theta}\beta^m)^{-1}[\nabla h(z^{k,m}) - \nabla h(x^k)], \widetilde{x} - z^{k,m}\rangle + \varphi(\widetilde{x}) - \varphi(z^{k,m}) \geq 0.$$

Soit $\widetilde{g}$ un sous-gradient quelconque de φ en $\widetilde{x}$. Alors, par définition de $\widetilde{g}$, la relation ci-dessus, (H1) et (H2), nous avons

$$\begin{aligned}
\widetilde{\theta}(\|F(x^k)\| + \|\widetilde{g}\|)\|\widetilde{x} - z^{k,m}\| &\geq (\widetilde{\theta}\beta^m)(\langle F(x^k), \widetilde{x} - z^{k,m}\rangle + \langle \widetilde{g}, \widetilde{x} - z^{k,m}\rangle) \\
&\geq (\widetilde{\theta}\beta^m)(\langle F(x^k), \widetilde{x} - z^{k,m}\rangle + \varphi(\widetilde{x}) - \varphi(z^{k,m})) \\
&\geq \langle \nabla h(z^{k,m}) - \nabla h(\widetilde{x}), z^{k,m} - \widetilde{x}\rangle \\
&\quad + \langle \nabla h(\widetilde{x}) - \nabla h(x^k), z^{k,m} - \widetilde{x}\rangle \\
&\geq \tau'\|z^{k,m} - \widetilde{x}\|^2 - \|\nabla h(\widetilde{x}) - \nabla h(x^k)\|\|z^{k,m} - \widetilde{x}\| \\
&\geq \tau'\|z^{k,m} - \widetilde{x}\|^2 - \tau''\|x^k - \widetilde{x}\|\|z^{k,m} - \widetilde{x}\|
\end{aligned}$$

i.e.

$$\|\widetilde{x} - z^{k,m}\| \leq \|x^k - \widetilde{x}\|\frac{\tau''}{\tau'} + \frac{\widetilde{\theta}}{\tau'}(\|F(x^k)\| + \|\widetilde{g}\|).$$

Il suit que

$$\begin{aligned}
\|x^k - z^{k,m}\| &\leq \|x^k - \widetilde{x}\| + \|\widetilde{x} - z^{k,m}\| \\
&\leq \|x^k - \widetilde{x}\|(1 + \tau''/\tau') + \widetilde{\theta}(\|F(x^k)\| + \|\widetilde{g}\|)/\tau' \\
&\equiv \lambda_k.
\end{aligned}$$

Soit L_k la constante de Lipschitz de F sur l'ensemble

$$\{x \in V | \|x - x^k\| \leq \lambda_k\}.$$

Alors, en utilisant les propriétés de h nous avons

$$\begin{array}{rl} & \langle F(x^k) - F(z^{k,m}), x^k - z^{k,m}\rangle \\ \leq & \|F(x^k) - F(z^{k,m})\|\|x^k - z^{k,m}\| \\ \leq & L_k\|x^k - z^{k,m}\|^2 \\ \leq & L_k\langle \nabla h(z^{k,m}) - \nabla h(x^k), z^{k,m} - x^k\rangle/\tau' \\ \leq & (1-\alpha)(\widetilde{\theta}\beta^m)^{-1}\langle \nabla h(z^{k,m}) - \nabla h(x^k), z^{k,m} - x^k\rangle \end{array}$$

où $(1-\alpha)(\widetilde{\theta}\beta^m)^{-1} \geq L_k/\tau'$ i.e. $\widetilde{\theta}\beta^m = \theta_k \leq (1-\alpha)\tau'/L_k$.

En prenant en compte (3.2) nous avons

$$\widetilde{\theta}\beta^{m-1} = \frac{\theta_k}{\beta} \geq \frac{(1-\alpha)\tau'}{L_k}.$$

Nous obtenons alors

$$\theta_k \geq \min\{\beta(1-\alpha)\tau'/L_k, \widetilde{\theta}\} > 0 \tag{3.11}$$

et par conséquent l'assertion (1) est vraie.

2. Puisque y^k est solution du problème (3.3)

$$\begin{array}{ll} \min & f(y) \\ \text{s.c.} & y \in C \quad \text{où} \quad f(y) = \langle F(x^k) - (\widetilde{\theta}\beta^m)^{-1}\nabla h(x^k), y\rangle + (\widetilde{\theta}\beta^m)^{-1}h(y) + \varphi(y) \end{array}$$

nous avons

$$\langle \nabla f(y^k), x - y^k\rangle \geq 0 \quad \forall x \in C$$

qui, en remplacant $f(y^k)$ par son expression, est équivalent à

$$\langle F(x^k) - (\widetilde{\theta}\beta^m)^{-1}\nabla h(x^k), y\rangle + (\widetilde{\theta}\beta^m)^{-1}\nabla h(y^k) + \nabla\varphi(y^k), x - y^k \geq 0.$$

$(\widetilde{\theta}\beta^m)^{-1}h + \varphi$ étant fortement convexe (de module $\tau' > 0$), nous avons alors

$$\langle(\widetilde{\theta}\beta^m)^{-1}\nabla h(y^k) + \nabla\varphi(y^k) - (\widetilde{\theta}\beta^m)^{-1}\nabla h(x^k) - \nabla\varphi(x^k), y^k - x^k\rangle \geq \tau'\|y^k - x^k\|^2.$$

Ainsi

$$\begin{array}{rcl} -\langle F(x^k) + \nabla\varphi(x^k), y^k - x^k\rangle & \geq & \langle(\widetilde{\theta}\beta^m)^{-1}\nabla h(y^k) + \nabla\varphi(y^k) \\ & & -(\widetilde{\theta}\beta^m)^{-1}\nabla h(x^k) - \nabla\varphi(x^k), y^k - x^k\rangle \\ & \geq & \tau'\|y^k - x^k\|^2. \end{array}$$

D'où

$$\|F(x^k) + \nabla\varphi(x^k)\|\|y^k - x^k\| \geq \tau'\|y^k - x^k\|^2$$

et

$$\|y^k - \underbrace{x^k}_{\text{bornée}}\| \leq \underbrace{\frac{1}{\tau'}\|F(x^k) - \nabla\varphi(x^k)\|}_{\text{borné}}.$$

Donc $\{y^k\}$ est bornée.

Soit

$$S = conv\left\{\overline{\bigcup_k B(x^k, \lambda_k)}\right\} \subseteq \underbrace{V}_{\text{fermé convexe}}.$$

La suite $\{x^k\}$ étant bornée et puisque $\lambda_k \leq \lambda' < +\infty$, nous avons que l'ensemble S est compact et convexe.

L'opérateur F étant localement Lipschitz continu sur C par hypothèse et S étant compact et convexe sur C, nous avons, par la Théorème 3.1.2 de [2] I (p. 174), que F est Lipschitz sur S.

Il suit qu'il existe une constante $L > 0$ telle que pour tout $x, y \in S$

$$\|F(x) - F(y)\| \leq L\|x - y\|.$$

Nous avons alors

$$L_k \leq L < +\infty \quad \forall k.$$

Par conséquent, (3.11) implique (3.10) où $\theta' = \min\{\beta(1-\alpha)\tau'/L', \widetilde{\theta}\}$ et la preuve est complète.

□

Lemme 3.2.2 *1. Si $x^k = y^k$, alors x^k résout (GVIP) (3.3).*

2. Si $x^k \neq y^k$, alors

$$\begin{aligned} \langle g^k, x^k - y^k\rangle &\geq (\alpha/\theta_k)\langle\nabla h(y^k) - \nabla h(x^k), y^k - x^k\rangle \\ &\geq (\alpha\tau'/\theta_k)\|x^k - y^k\|^2. \end{aligned} \tag{3.12}$$

Preuve.

1. Nous remarquons tout d'abord que $\theta_k > 0$ dû au Lemme 3.2.1 (1). Par conséquent, si $x^k = y^k$, alors par le Lemme 3.1.1 (1) et (3.8) nous avons

$$\begin{aligned} \langle F(x^k), y - x^k\rangle + \varphi(y) - \varphi(x^k) &= \langle F(x^k), y - x^k\rangle + \varphi(y) - \varphi(x^k) \\ &\quad + \theta_k^{-1}\langle\nabla h(x^k) - \nabla h(x^k), y - x^k\rangle \\ &\geq 0 \end{aligned}$$

pour chaque $y \in C$, i.e., x^k résout $(GVIP)$ (3.3) et la partie (1) est vraie.

2. Dans le cas où $x^k \neq y^k$, par (3.4), (3.2) et la Proposition B.0.1, nous avons

$$\begin{aligned}\langle g^k, x^k - y^k\rangle &= \langle F(y^k) - F(x^k) - \theta_k^{-1}[\nabla h(y^k) - \nabla h(x^k)], x^k - y^k\rangle \\ &\geq ((1-\alpha)/\theta_k)\langle \nabla h(y^k) - \nabla h(x^k), x^k - y^k\rangle \\ &\quad -(1/\theta_k)\langle \nabla h(y^k) - \nabla h(x^k), x^k - y^k\rangle \\ &= (\alpha/\theta_k)\langle \nabla h(y^k) - \nabla h(x^k), y^k - x^k\rangle \\ &\geq (\alpha\tau'/\theta_k)\|x^k - y^k\|^2\end{aligned}$$

i.e., la partie (2) est aussi vraie.

□

Nous remarquons que $x^k \neq y^k$ implique $g^k \neq 0$ dû à (3.12), donc que le Pas 2 est bien défini.

Lemme 3.2.3 *1. Si x^k résout $(GVIP)$ (3.3), alors*

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - \gamma_k(2-\gamma_k)(\sigma_k\|g^k\|)^2.$$

2. $\{x^k\}$ est bornée.

3. $\sum_{k=0}^{+\infty} \gamma_k(2-\gamma_k)(\sigma_k\|g^k\|)^2 < +\infty$.

4. Pour chaque valeur d'adhérence x^* de $\{x^k\}$ telle que x^* résolve $(GVIP)$ (3.3), nous avons

$$\lim_{k\to+\infty} x^k = x^*.$$

Preuve.
Soit $x^* \in C^*$ une solution de $(GVIP)$. Alors, par monotonicité de F, nous avons

$$\langle F(y^k), y^k - x^*\rangle + \varphi(y^k) - \varphi(x^*) \geq \langle F(x^*), y^k - x^*\rangle + \varphi(y^k) - \varphi(x^*) \geq 0.$$

En tenant compte du Lemme 3.1.1 (1), en appliquant (3.8) avec $y = x^*$, $\overline{z} = z^{k,m} = y^k$, nous avons

$$\langle -F(x^k) - \theta_k^{-1}[\nabla h(y^k) - \nabla h(x^k)], y^k - x^*\rangle + \varphi(x^*) - \varphi(y^k) \geq 0.$$

En additionnant les deux inégalités au-dessus, nous avons, par (3.4), que $\langle g^k, y^k - x^*\rangle \geq 0$ d'où par (3.5)

$$\langle g^k, x^k - x^*\rangle = \langle g^k, x^k - y^k\rangle + \langle g^k, y^k - x^*\rangle \geq \sigma_k\|g^k\|^2 \geq 0.$$

Cette relation prouve la partie (1) dû à (3.6), (3.5) et (3.1). Les parties (2)-(4) suivent directement de (1).

□

Nous sommes maintenant en position d'établir un résultat de convergence pour la Méthode 3.1.1.

Théorème 3.2.1 *Soit $\{x^k\}$ une suite construite par la Méthode 3.1.1. Si la méthode finit à la k-ième itération, alors x^k résout (GVIP) (3.3). Autrement, si x^k est infinie, alors x^k converge vers une solution de (GVIP) (3.3).*

Preuve.
La première assertion suit du Lemme 3.2.2 (1). Soit $\{x^k\}$ une suite infinie. Alors, par les Lemmes 3.2.3 (2) et 3.2.1 (2), $\{x^k\}$, $\{F(x^k)\}$, $\{y^k\}$, et $\{F(y^k)\}$ sont bornées. Ensuite, par (3.4), (3.5), (3.12) et (H1), nous avons

$$\begin{aligned} \sigma_k\|g^k\| &= \langle g^k, x^k - y^k\rangle / \|F(y^k) - F(x^k) - \theta_k^{-1}[\nabla h(y^k) - \nabla h(x^k)]\| \\ &\geq \alpha\tau'/\theta_k\|x^k - y^k\|^2/((L + \tau''/\theta_k)\|x^k - y^k\|) \\ &\geq \alpha\tau'\|x^k - y^k\|/(\widetilde{\theta}L + \tau'') \end{aligned} \tag{3.13}$$

où L est la constante de Lipschitz de F. D'un autre côté, (3.1) et le Lemme 3.2.3 (3) impliquent que

$$\liminf_{k\to+\infty}(\sigma_k\|g^k\|) = 0.$$

Il suit maintenant que

$$\liminf_{k\to+\infty}\|x^k - y^k\| = 0. \tag{3.14}$$

Puisque $\{x^k\}$ et $\{y^k\}$ sont bornées, elles ont des valeurs d'adhérence. Sans perdre de généralité nous pouvons conclure qu'il existe des sous-suites $\{x^{k_s}\}$ et $\{y^{k_s}\}$ telles que

$$\lim_{s\to+\infty} x^{k_s} = \lim_{s\to+\infty} y^{k_s} = x^* \in C.$$

Soit $y \in C$. En prenant en compte le Lemme 3.1.1 (1), (3.8) avec $\overline{z} = y^{k_s}$, (3.10), et les propriétés de h, nous avons

$$\begin{aligned} \langle F(x^{k_s}), y - y^{k_s}\rangle + \varphi(y) - \varphi(y^{k_s}) &\geq -\langle\nabla h(y^{k_s}) - \nabla h(x^{k_s}), y - y^{k_s}\rangle/\theta_{k_s} \\ &\geq -\|\nabla h(x^{k_s}) - \nabla h(y^{k_s})\|\|y - y^{k_s}\|/\theta_{k_s} \\ &\geq -\tau''\|x^{k_s} - y^{k_s}\|\|y - y^{k_s}\|/\theta'. \end{aligned}$$

En prenant la limite lorsque s tend vers $+\infty$ dans cette inégalité, nous avons

$$\langle F(x^*), y - x^*\rangle + \varphi(y) - \varphi(x^*) \geq 0$$

i.e. x^* résout $(GVIP)$. Le résultat désiré suit alors du Lemme 3.2.3 (4). □

3.3 Taux de convergence

En tenant compte du Lemme 3.2.2 et des propriétés de h, la valeur de $\langle \nabla h(y^k) - \nabla h(x^k), y^k - x^k \rangle$ peut être vue comme une borne d'erreur de la Méthode 3.1.1. Le résultat suivant est basé sur cette observation.

Théorème 3.3.1 *Soit $\{x^k\}$ une suite infine générée par la Méthode 3.1.1. Supposons que*

$$\gamma_k = \gamma \in (0,2), \quad k = 0, 1, \dots \tag{3.15}$$

Alors

$$\liminf_{k \to +\infty} (\langle \nabla h(y^k) - \nabla h(x^k), y^k - x^k \rangle \sqrt{k+1}) = 0 \tag{3.16}$$

Preuve.
Des Lemmes 3.2.3 (2) et 3.2.1 (2) nous avons que les suites $\{x^k\}$, $\{y^k\}$ et $\{F(y^k)\}$ sont bornées. En combinant (3.4), (3.5) et (H2) nous avons alors

$$\begin{aligned}
(\sigma_k \|g^k\|)^2 &= \langle g^k, x^k - y^k \rangle^2 / \|F(y^k) - F(x^k) - \theta_k^{-1}[\nabla h(y^k) - \nabla h(x^k)]\|^2 \\
&\geq (\alpha/\theta_k)^2 \langle \nabla h(y^k) - \nabla h(x^k), y^k - x^k \rangle^2 / ((L + \theta_k^{-1}\tau'')\|x^k - y^k\|)^2 \\
&= \alpha^2 \langle \nabla h(y^k) - \nabla h(x^k), y^k - x^k \rangle^2 / ((\theta_k L + \tau'')\|x^k - y^k\|)^2 \\
&\geq \alpha^2 \langle \nabla h(y^k) - \nabla h(x^k), y^k - x^k \rangle^2 / ((\widetilde{\theta} L + \tau'')\|x^k - y^k\|)^2 \\
&\geq \tau'\alpha^2 \langle \nabla h(y^k) - \nabla h(x^k), y^k - x^k \rangle / (\widetilde{\theta} L + \tau'')^2.
\end{aligned}$$

Par le Lemme 3.2.3 (3) et (3.15), nous avons

$$\sum_{k=0}^{+\infty} \gamma(2-\gamma)(\sigma_k \|g^k\|)^2 < +\infty.$$

Si nous supposons que (3.16) n'a pas lieu, alors il existe un nombre $\mu > 0$ tel que

$$\langle \nabla h(y^k) - \nabla h(x^k), y^k - x^k \rangle \geq \frac{\mu}{\sqrt{k+1}}$$

et l'inégalité au-dessus implique alors

$$+\infty > \left(\frac{\tau'\alpha^2\mu}{(\widetilde{\theta} L + \tau'')^2} \right)^2 \sum_{k=0}^{+\infty} \gamma(2-\gamma)$$

qui contredit (3.1). □

Selon le Corollaire B.0.1, $(GVIP)$ a une solution unique si F est fortement monotone. Nous montrons que notre méthode converge au moins linéairement sous cette hypothèse.

Théorème 3.3.2 *Supposons que F soit fortement monotone de constante κ et que (3.15) ait lieu. Si la Méthode 3.1.1 génère une suite infinie $\{x^k\}$, alors $\{x^k\}$ converge vers une solution du problème $(GVIP)$ avec un taux linéaire.*

Preuve.

– Nous montrons tout d'abord qu'il existe un nombre $\mu > 0$ tel que

$$\|x^k - y^k\|^2 \geq \mu \|x^k - x^*\|^2 \tag{3.17}$$

où x^* est une solution unique du problème $(GVIP)$. En effet, de $(GVIP)$ avec $y = y^k$ nous avons

$$\langle F(x^*), y^k - x^* \rangle + \varphi(y^k) - \varphi(x^*) \geq 0.$$

En tenant compte du Lemme 3.1.1 (1), et en appliquant (3.8) avec $\overline{z} = y^k$ et $y = x^*$ nous avons

$$\langle F(x^k) + \theta_k^{-1}[\nabla h(y^k) - \nabla h(x^k)], x^* - y^k \rangle + \varphi(x^*) - \varphi(y^k) \geq 0.$$

En additionnant ces deux inégalités, nous avons

$$\begin{aligned} \langle F(x^*) - F(x^k), y^k - x^k \rangle + \langle F(x^*) - F(x^k), x^k - x^* \rangle \\ + \theta_k^{-1} \langle \nabla h(y^k) - \nabla h(x^k), x^* - y^k \rangle \geq 0 \end{aligned}$$

i.e.,

$$\begin{aligned} \langle F(x^k) - F(x^*), x^k - x^* \rangle \ \leq\ & \langle F(x^*) - F(x^k), y^k - x^k \rangle \\ & + \theta_k^{-1} \langle \nabla h(y^k) - \nabla h(x^k), x^* - y^k \rangle. \end{aligned}$$

En prenant en compte les Lemmes 3.2.3 (2) et 3.2.1 (2) et les propriétés de h et de F, nous avons

$$\begin{aligned} \kappa \|x^k - x^*\|^2 &\leq \langle F(x^k) - F(x^*), x^k - x^* \rangle \\ &\leq \langle F(x^*) - F(x^k), y^k - x^k \rangle \\ &\quad + \theta_k^{-1} \langle \nabla h(y^k) - \nabla h(x^k), x^* - y^k \rangle \\ &\leq (L + \tau''/\theta_k) \|x^* - x^k\| \|y^k - x^k\| \\ &\leq (L + \tau''/\theta') \|x^* - x^k\| \|y^k - x^k\| \end{aligned}$$

où L est la constante de Lipschitz de F, i.e., (3.17) a lieu avec $\mu = (\kappa/(L + \tau''/\theta'))^2$.

– En utilisant (3.13), nous obtenons alors

$$\begin{aligned} (\sigma_k \|g^k\|)^2 &\geq (\alpha\tau' \|x^k - y^k\| / (\widetilde{\theta} L + \tau''))^2 \\ &\geq (\alpha\tau')^2 \|x^k - x^*\|^2 \mu / (\widetilde{\theta} L + \tau'')^2 \\ &\equiv \mu' \|x^k - x^*\|^2. \end{aligned}$$

Cette inégalité, avec (3.15) et le Lemme 3.2.3 (1), donne

$$\begin{aligned}\|x^{k+1} - x^*\|^2 &\leq \|x^k - x^*\|^2 - \gamma(2-\gamma)\mu'\|x^k - x^*\|^2 \\ &= (1 - \gamma(2-\gamma)\mu')\|x^k - x^*\|^2\end{aligned}$$

et le résultat suit. □

Nous donnons maintenant des conditions qui assurent la terminaison finie de la Méthode 3.1.1. Nous considérons l'hypothèse suivante.

(H3) Il existe un nombre $\mu' > 0$ tel que, pour chaque point $x \in C$, l'inégalité suivante a lieu :

$$\langle F(x), x - \pi_{C^*}(x)\rangle + \varphi(x) - \varphi(\pi_{C^*}(x)) \geq \mu'\|x - \pi_{C^*}(x)\|.$$

Théorème 3.3.3 *Soit $\{x^k\}$ une suite construite par la Méthode 3.1.1. Supposons que (H3) ait lieu. Alors la méthode termine avec une solution.*

Preuve.
Nous supposons par contradiction que la suite $\{x^k\}$ est infinie.

Alors, suivant la preuve du Théorème 3.2.1, nous voyons que (3.14) a lieu, de plus, les suites $\{x^k\}$ et $\{y^k\}$ sont bornées.

D'un autre côté, par (H1)-(H3), les Lemmes 3.1.1 (3.8) et 3.2.1 (2) il suit que

$$\begin{aligned}\mu'\|y^k - \pi_{C^*}(y^k)\| &\leq \langle F(y^k), y^k - \pi_{C^*}(y^k)\rangle + \varphi(y^k) - \varphi(\pi_{C^*}(y^k)) \\ &= \langle F(y^k) - F(x^k) - \theta_k^{-1}(\nabla h(y^k) - \nabla h(x^k)), y^k - \pi_{C^*}(y^k)\rangle \\ &\quad + \langle F(x^k) + \theta_k^{-1}(\nabla h(y^k) - \nabla h(x^k)), y^k \pi_{C^*}(y^k)\rangle \\ &\quad + \varphi(y^k) - \varphi(\pi_{C^*}(y^k)) \\ &\leq \langle F(y^k) - F(x^k) - \theta_k^{-1}(\nabla h(y^k) - \nabla h(x^k)), y^k - \pi_{C^*}(y^k)\rangle \\ &\leq (L\|y^k - x^k\| + \theta_k^{-1}\tau''\|y^k - x^k\|)\|y^k - \pi_{C^*}(y^k)\| \\ &\leq (L + \tau''/\theta')\|y^k - x^k\|\|y^k - \pi_{C^*}(y^k)\|\end{aligned}$$

où L est la constante de Lipschitz pour F. Il suit que

$$\|y^k - x^k\| \geq \mu'/(L + \tau''/\theta').$$

Nous avons donc une contradiction et la thèse est démontrée. □

Chapitre 4

Méthode de relaxation combinée pour le problème $(MVIP)$

Nous considérons à présent une méthode pour résoudre le problème $(MVIP)$. Celle-ci sera différente de celle établie pour le problème $(GVIP)$ car nous devons prendre en compte la situation du pas nul dans le cas multivoque. Avant de décrire la méthode, nous intoduisons les définitions suivantes ainsi que les hypothèses générales du chapitre.

Définition 4.0.1 *Soient W et V deux ensembles convexes dans $\mathbb{R}^n$ tels que $W \subseteq V$. Soit $P : V \to \Pi(\mathbb{R}^n)$ un opérateur multivoque.*
L'opérateur P est dit :

- *semi-continu supérieurement (u.s.c.) en x si $\forall \epsilon > 0$ il existe un voisinage $N(x)$ de x tel que*
$$\forall y \in N(x), \quad F(y) \subseteq F(x) + B(0, \epsilon);$$
- *un K-mapping sur W, s'il est u.s.c. sur W et a des valeurs compactes non vides et convexes.*

Les hypothèses générales (HG) de ce chapitre sont les suivantes :

1. C est un sous-ensemble de $\mathbb{R}^n$ défini par
$$C = \{x \in \mathbb{R}^n | h(x) \leq 0\} \tag{4.1}$$
où $h : \mathbb{R}^n \to \mathbb{R}$ est une fonction convexe mais non nécessairement différentiable ;
2. la condition de Slater est satisfaite, i.e., il existe un point $\overline{x}$ tel que $h(\overline{x}) < 0$;
3. $F : C \to \Pi(\mathbb{R}^n)$ est un K-mapping ;
4. $C^d \neq \emptyset$.

Remarque
L'ensemble admissible est habituellement défini par

$$C = \{x \in \mathbb{R}^n | h_i(x) \leq 0 \quad i = 1, \ldots, m\}$$

où $h_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$ sont des fonctions convexes. Mais, en posant $h(x) = \max_{i=1,\ldots,m} h_i(x)$ nous pouvons facilement réduire ce cas à l'initial (4.1) sans perdre de généralité.

4.1 Description de la méthode

Nous définissons l'opérateur $Q : \mathbb{R}^n \to \Pi(\mathbb{R}^n)$ par

$$Q(x) = \begin{cases} F(x) & \text{si} \quad h(x) \leq 0, \\ \partial h(x) & \text{si} \quad h(x) > 0. \end{cases} \tag{4.2}$$

Méthode 4.1.1 *Choisir un point $x^0 \in C$, une suite $\{\gamma_j\}$ telle que*

$$\gamma_j \in [0,2],\ k = 0,1,\ldots; \quad \sum_{j=0}^{+\infty} \gamma_j(2-\gamma_j) = +\infty\,; \tag{4.3}$$

et deux suites positives bornées $\{\epsilon_l\}$ et $\{\eta_k\}$. Choisir un nombre $\theta \in (0,1)$ et une suite d'opérateurs $\{P_k\}$, où $P_k \in \mathcal{F}(C)$ pour $k = 0,1,\ldots$. Poser $k=0$, $l=1$, $u^0 = x^0$, $j(0)=0$, $k(0)=0$, $t=0$ et $K_l = \{0\}$.

Pas 1 (Procédure auxilliaire $\widetilde{D}_k$)

Pas 1.1 : Choisir q^0 dans $Q(x^k)$, poser $i=0$, $p^i = q^i$, $I_{k,l} = \{0\}$, $t = t+1$ et $w^0 = x^k$.

Pas 1.2 (Pas nul) : Si

$$\|p^i\| \leq \eta_l \tag{4.4}$$

poser $k(l) = k$, $u^l = x^{k+1}$, $j(k+1) = j(k)$, $k = k+1$, $l = l+1$ et $K_l = \{k\}$ et retour au Pas 1.

Pas 1.3 (Pas de descente) : Poser $w^{i+1} = x^k - \epsilon_l p^i/\|p^i\|$, choisir $q^{i+1} \in Q(w^{i+1})$, poser $I_{k,l} = I_{k,l}\bigcup\{i+1\}$ et $t = t+1$. Si

$$\langle q^{i+1}, p^i\rangle > \theta\|p^i\|^2 \tag{4.5}$$

alors poser $y^k = w^{i+1}$, $g^k = q^{i+1}$ et $j(k+1) = j(k)+1$ et aller au Pas 2.

Pas 1.4 : Poser

$$p^{i+1} = Nrconv\{p^i, q^{i+1}\} \tag{4.6}$$

$i = i+1$ et retour au Pas 1.2.

Pas 2 (Itération principale) *Poser $\sigma_k = \langle g^k, x^k - y^k\rangle/\|g^k\|^2$,*

$$x^{k+1} = P_k(x^k - \gamma_{j(k)}\sigma_k g^k)$$

$k = k+1$, $K_l = K_l\bigcup\{k\}$ et retour au Pas 1.

Selon cette description, à chaque itération, la procédure auxiliaire au Pas 1 est appliquée pour trouver la direction g^k.

Dans le cas d'un pas nul, les tolérances ϵ_l et η_l diminuent puisque le point x^k approxime une solution avec ϵ_l et η_l (cfr. Lemme 4.3.2 (2)). Par conséquent, la variable l est un compteur pour les pas nuls et la variable $j(.)$ est un compteur pour les pas de descente.

Dans le cas d'un pas de descente, nous devons avoir $\sigma_k > 0$ (cfr. Lemme 4.3.2 (3)). Donc, le point $\overline{x}^{k+1} = x^k - \gamma\sigma_k g^k$ est la projection du point x^k sur l'hyperplan

$$H_k(\gamma) = \{y \in \mathbb{R}^n | \langle g^k, y - x^k \rangle = -\gamma\sigma_k \|g^k\|^2\}.$$

Clairement, $H_k(1)$ sépare x^k et C^d. D'où, la distance de $\overline{x}^{k+1}$ à chaque point de C^d ne peut augmenter lorsque $\gamma \in (0, 2)$ et celle de x^{k+1} fait de même dû aux propriétés de P_k.

Remarque

Les indices t et $k(l)$, les ensembles d'indices $I_{k,l}$ et K_l, et la suite $\{u^l\}$ sont introduits seulement par simplicité mais peuvent être éliminés. Nous appellerons une augmentation de l'indice t un pas intérieur, donc le nombre de pas intérieurs donne le nombre de calculs d'éléments de $Q(\cdot)$ aux points correspondants.

4.2 Propriétés des opérateurs auxiliaires

Nous définissons l'opérateur $P : \mathbb{R}^n \to \Pi(\mathbb{R}^n)$ par

$$P(x) = \begin{cases} F(x) & \text{si} \quad h(x) < 0, \\ conv\{G(x) \bigcup \partial h(x)\} & \text{si} \quad h(x) = 0, \\ \partial h(x) & \text{si} \quad h(x) > 0. \end{cases} \tag{4.7}$$

Nous obtenons à présent certaines propriétés des opérateurs Q et P, que nous avons l'intention d'utiliser pour prouver la convergence de notre méthode. En particulier, nous voulons montrer que $(MVIP)$ est équivalent à l'inclusion multivoque

$$0 \in P(x^*). \tag{4.8}$$

Nous notons par S^* l'ensemble des solutions du problème (4.8) .

Nous établissons tout d'abord les propriétés de continuité pour Q et P.

Proposition 4.2.1 *Les opérateurs Q et P, définis par (4.2) et (4.7) sont des K-mappings.*

Nous notons par $coneC$ l'enveloppe conique de C, i.e.,

$$coneC = \{q \in \mathbb{R}^n | q = \lambda x, \lambda \geq 0, x \in C\}.$$

Lemme 4.2.1 *Soient W et Y des ensembles compactes convexes dans $\mathbb{R}^n$, $0 \in Y$. Alors la relation $0 \in conv\{W \bigcup Y\}$ est équivalente à $0 \in W + coneY$.*

Lemme 4.2.2 *Si $x \in C$, alors*

$$N(C, x) = \begin{cases} 0 & si \quad h(x) < 0, \\ cone\partial h(x) & si \quad h(x) = 0. \end{cases}$$

Théorème 4.2.1 *Nous avons $C^* = S^*$.*

Preuve.
Il suit de la condition de Slater que $0 \notin \partial h(x)$ si $h(x) \geq 0$. Donc, si (4.8) a lieu, alors $x^* \in C$. En posant $W = F(x^*)$ et $Y = \partial h(x^*)$ dans le Lemme 4.2.1, nous voyons que (4.8) est équivalent à

$$x^* \in C, \quad 0 \in \begin{cases} F(x^*) & \text{si} \quad h(x^*) < 0, \\ F(x^*) + cone\partial h(x^*) & \text{si} \quad h(x^*) = 0. \end{cases}$$

Il suit à présent du Lemme 4.2.2 que (4.8) est équivalent à $(MVIP)$. $\square$

Nous établissons à présent un résultat de convergence pour la Méthode 4.1.1.

Théorème 4.3.1 *Soit $\{x^k\}$ une suite générée par la Méthode 4.1.1 et soient $\{\epsilon_l\}$ et $\{\eta_l\}$ satisfaisant les relations suivantes :*

$$\{\epsilon_l\} \searrow 0, \{\eta_l\} \searrow 0. \tag{4.13}$$

Alors

1. *le nombre de pas intérieurs à chaque itération est fini ;*
2. *il existe une valeur d'adhérence x^* de $\{x^k\}$ qui se trouve dans C^* ;*
3. *si, en plus,*
$$C^* = C^d, \tag{4.14}$$
nous avons
$$\lim_{k\to+\infty} x^k = x^* \in C^*.$$

Preuve.

1. Suit des Lemmes 4.3.2 (1) et 4.3.3 (2).

2. Si $l \leq s < +\infty$ alors il existe un nombre $\overline{k}$ tel que toutes les itérations de la procédure auxiliaire se terminent au Pas 1.3 lorsque $k \geq \overline{k}$.

   Par le Lemme 4.3.3 (2), $\{y^k\}$ est bornée, donc $\{g^k\}$ l'est aussi, i.e.,
$$\|g^k\| \leq C_2 < +\infty, \quad k = 0, 1, \ldots$$
   En prenant en compte le Lemme 4.3.2 (3), nous avons alors
$$\sigma_k\|g^k\| = \frac{\langle g^k, x^k - y^k\rangle}{\|g^k\|} > \frac{\theta\epsilon_l\eta_l}{C_2} \equiv \widetilde{\sigma} > 0$$
   donc, par le Lemme 4.3.3 (4), la suite $\{u^l\}$ est infinie. De plus, $\{u^l\}$ est bornée dû au Lemme 4.3.3 (2), par conséquent, il existe une sous-suite de $\{u^l\}$ qui converge vers un certain point $x^* \in C$. Sans perdre de généralité, nous supposons que $\lim_{k\to+\infty} u^l = x^*$.

   D'un autre côté, selon le Lemme 4.3.2 (2) et la construction de la Méthode 4.1.1, il existe un élément $p^{(l)}$ tel que
$$\|p^{(l)}\| \leq \eta_l, \quad p^{(l)} \in conv \bigcup_{\|y-u^l\|\leq\epsilon_l} P(y)$$
   pour $l = 1, 2, \ldots$

Puisque P est u.s.c. dû à la Proposition 4.2.1, ces propriétés avec (4.13) nous permettent de dire que $x^* \in S^*$ ou, de manière équivalente, $x^* \in C^*$ dû au Théorème 4.2.1.

3. Suit de (4.14) et du Lemme 4.3.3 (5). $\square$

4.4 Estimations de la complexité

Dans cette section, nous donnons certaines estimations de la complexité pour la Méthode 4.1.1. Comme la Méthode 4.1.1 a une structure à deux niveaux avec chaque itération contenant un nombre fini de pas intérieurs, elle est plus adaptée pour obtenir ces estimations de la complexité.

Nous considérons les hypothèses additionnelles (HA) suivantes :

1. $$\gamma_k = \gamma \in (0, 2) \quad \text{pour } k = 0, 1, \ldots ; \tag{4.15}$$

2. il existe $x^* \in C^*$ tel que pour chaque $x \in C$ et pour chaque $g \in F(x)$,

$$\langle g, x - x^* \rangle \geq \mu \|x - x^*\|^{\epsilon} \tag{4.16}$$

où $\epsilon > 0$ et $\mu > 0$.

Il est facile de voir que (4.16) généralise légèrement $(DMVIP)$ dans le cas où $\mu = 0$. D'un autre côté, si F est fortement monotone de constante μ, alors (4.16) a lieu avec $\epsilon = 2$. Nous remarquons aussi que la condition (4.16) implique que $C^* = \{x^*\}$, i.e., C^* contient au plus un point.

Dans cette section, nous utilisons la distance de x^* comme une fonction de précision pour notre méthode. Plus précisément, avec x^0 un point de départ donné et un nombre $\delta > 0$, nous définissons la complexité de la méthode, notée par $N(\delta)$, comme le nombre total de pas intérieurs t qui assure de trouver un point $\overline{x} \in C$ tel que

$$\|\overline{x} - x^*\| \leq \delta.$$

Rappelons que chaque pas intérieur inclut un calcul d'un élément de $Q(.)$ au point correspondant. Donc, puisque la dépense en calcul par pas intérieur peut facilement être évalué pour chaque problème spécifique sous examination, cette estimation donne en fait la quantité totale de travail.

Nous procédons maintenant à obtenir une borne supérieure pour $N(\delta)$. Nous établissons tout d'abord plusieurs propriétés auxiliaires.

Lemme 4.4.1 *Soit $\{x^k\}$ une suite générée par la Méthode 4.1.1 et supposons que les hypothèses (HA) soient satisfaites. Alors*

$$\|g^k\| \leq C' < +\infty \quad et \quad D_{k,l} \leq C' < +\infty. \tag{4.17}$$

Dans ce qui suit, par simplicité, nous posons $D_l = D_{k(l),l}$. Nous fixons un point $\overline{u} \in C$ tel que $h(\overline{u}) < 0$.

Lemme 4.4.2 *Soit $\{x^k\}$ une suite générée par la Méthode 4.1.1 et supposons que les hypothèses (HA) soient satisfaites. Pour chaque $i = 1, 2, \ldots$, il existe un nombre $\tau_l \in [0, 1]$ tel que nous ayons :*

1. *si (4.16) a lieu avec $\epsilon = 1$, alors*

$$\eta_l \|x^* - u^l\| \geq \tau_l \mu \|x^* - u^l\| - \epsilon_l(\tau_l \mu + D_l); \tag{4.18}$$

2. *si F est monotone sur C, alors*

$$\eta_l \|\overline{u} - u^l\| + \epsilon_l D_l \geq \tau_l(\langle \overline{g}, u^l - \overline{u}\rangle - \epsilon_l \|\overline{g}\|) - (1 - \eta_l)h(\overline{u}) \tag{4.19}$$

où $\overline{g} \in F(\overline{u})$.

Preuve.
Nous fixons $l = 1, 2, \ldots$. Nous considérons la $k(l)$-ième itération de la Méthode 4.1.1. Par (4.9), pour chaque s, nous avons

$$p^s = \sum_{i=0}^{s} \beta_i q^i, \quad \sum_{i=0}^{s} \beta_i = 1, \quad \beta_i \geq 0, \quad \text{pour } i = 0, \ldots, s. \tag{4.20}$$

Par définition, à la k-ième itération, la procédure auxiliaire produit un pas nul, i.e., il existe un nombre $\overline{s}$ tel que

$$\|p^{\overline{s}}\| \leq \eta_l \tag{4.21}$$

(cfr (4.4)). Par simplicité, nous omettons les indices de l'ensemble $I_{k(l),l}$ dans cette preuve, i.e., soit

$$I = \{0, \ldots, \overline{s}\}, \quad \overline{I} = \{i \in I | w^i \in C\}, \quad \widetilde{I} = \{i \in I | w^i \notin C\}.$$

Alors

$$q^i \in \begin{cases} F(w^i) & \text{si} \quad i \in \overline{I}, \\ \partial h(w^i) & \text{si} \quad i \in \widetilde{I}. \end{cases} \tag{4.22}$$

Poser

$$\tau_l = \sum_{i \in \overline{I}} \beta_i. \tag{4.23}$$

Il est évident que $\tau_l \in [0, 1]$. Nous observons aussi que

$$\|w^i - u^l\| \leq \epsilon_l \quad \text{pour } i \in I. \tag{4.24}$$

1. Dans ce cas, nous avons en tenant compte de (4.22) et de (4.16),

$$\langle q^i, w^i - x^* \rangle \geq \mu \|w^i - x^*\| \quad \text{pour } i \in \overline{I}$$

et

$$\langle q^i, w^i - x^* \rangle \geq h(w^i) - h(x^*) \geq 0 \quad \text{pour } i \in \widetilde{I}$$

où $\{x^*\} = C^*$.

En appliquant (4.24), nous avons

$$\begin{aligned} \langle q^i, u^l - x^* \rangle &= \langle q^i, u^l - w^i \rangle + \langle q^i, w^i - x^* \rangle \\ &\geq \langle q^i, u^l - w^i \rangle + \mu \|w^i - x^*\| \\ &\geq -\|q^i\| \, \|u^l - w^i\| + \mu(\|x^* - u^l\| - \|u^l - w^i\|) \\ &\geq -\epsilon_l \|q^i\| + \mu(\|x^* - u^l\| - \epsilon_l) \end{aligned}$$

pour $i \in \overline{I}$ et

$$\langle q^i, u^l - x^* \rangle \geq \langle q^i, u^l - w^i \rangle \geq -\epsilon_l \|q^i\|$$

pour $i \in \widetilde{I}$.

En sommant ces deux inégalités multipliées par β_i, sur $i \in I$ et en prenant en compte (4.21), (4.11), (4.20) et (4.23), nous obtenons

$$\begin{aligned} \eta_l \|u^l - x^*\| &\geq \langle p^{\overline{s}}, u^l - x^* \rangle \\ &\geq \sum_{i \in \overline{I}} \beta_i [-\epsilon_l \|q^i\| + \mu(\|x^* - u^l\| - \epsilon_l)] - \sum_{i \in \widetilde{I}} \beta_i \epsilon_l \|q^i\| \\ &\geq \tau_l \mu(\|x^* - u^l\| - \epsilon_l)) + \sum_{i \in I} \beta_i(-\epsilon_l \|q^i\|) \\ &\geq \tau_l \mu(\|x^* - u^l\| - \epsilon_l)) - \epsilon_l D_l \end{aligned}$$

i.e. (4.18) a lieu.

2. Par monotonicité de F et par (4.22), pour chaque $\overline{g} \in F(\overline{u})$, nous avons

$$\langle q^i, w^i - \overline{u} \rangle \geq \langle \overline{g}, w^i - \overline{u} \rangle \quad \text{pour } i \in \overline{I}.$$

Aussi, par (4.22) et la propriété du sous-gradient,

$$\langle q^i, w^i - \overline{u} \rangle \geq h(w^i) - h(\overline{u}) \geq -h(\overline{u}) \quad \text{pour } i \in \widetilde{I}.$$

En appliquant (4.24), nous avons

$$\begin{aligned} \langle q^i, u^l - \overline{u} \rangle &= \langle q^i, u^l - w^i \rangle + \langle q^i, w^i - \overline{u} \rangle \\ &\geq \langle q^i, u^l - w^i \rangle + \langle \overline{g}, w^i - \overline{u} \rangle \\ &\geq -\epsilon_l \|q^i\| + \langle \overline{g}, u^l - \overline{u} \rangle + \langle \overline{g}, w^i - u^l \rangle \\ &\geq -\epsilon_l \|q^i\| + \langle \overline{g}, u^l - \overline{u} \rangle - \epsilon_l \|\overline{g}\| \end{aligned}$$

pour $i \in \overline{I}$ et, respectivement,

$$\langle q^i, u^l - \overline{u} \rangle \geq -\epsilon_l \|q^i\| - h(\overline{u})$$

pour $i \in \widetilde{I}$.

En sommant ces inégalités multipliées par β_i, sur $i \in I$ et en prenant en compte (4.21), (4.11), (4.20) et (4.23), nous obtenons

$$\begin{aligned}
\eta_l \|\overline{u} - u^l\| &\geq \langle p^{\overline{s}}, u^l - \overline{u} \rangle \\
&\geq \sum_{i \in \overline{I}} \beta_i [-\epsilon_l \|q^i\| + \langle \overline{g}, u^l - \overline{u} \rangle - \epsilon_l \|\overline{g}\|] \\
&\quad + \sum_{i \in \widetilde{I}} \beta_i [-\epsilon_l \|q^i\| - h(\overline{u})] \\
&\geq -\epsilon_l D_l + \tau_l (\langle \overline{g}, u^l - \overline{u} \rangle - \epsilon_l \|\overline{g}\|) - (1 - \tau_l) h(\overline{u})
\end{aligned}$$

i.e. (4.19) a lieu. □

Nous montrons maintenant que la Méthode 4.1.1 atteint une estimation de complexité logarithmique. Nous notons par $l(\delta)$ la valeur maximale de l'indice l dans la Méthode 4.1.1 tel que

$$\|u^l - x^*\| \geq \delta.$$

En prenant en compte la définition de pas intérieur, nous pouvons conclure que la borne suivante est vraie :

$$N(\delta) \leq \sum_{l=1}^{l(\delta)+1} \sum_{k \in K_l} (|I_{k,l}| - 1). \tag{4.25}$$

Nous procédons maintenant à estimer le membre de droit de (4.25). Nous établissons tout d'abord une estimation de la complexité pour le cas multivoque.

Théorème 4.4.1 *Supposons que F soit monotone et que les hypothèses (HA) soient satisfaites avec $\epsilon = 1$. Soit $\{x^k\}$ une suite générée par la Méthode 4.1.1 où*

$$\epsilon_l = \nu^l \epsilon', \eta_l = \eta', l = 0, 1, \ldots; \quad \nu \in (0, 1). \tag{4.26}$$

Alors, il existe certaines constantes $\overline{\epsilon} > 0$ et $\overline{\eta}$ telles que

$$N(\delta) \leq B_l \nu^{-2} (\ln(B_0/\delta) / \ln \nu^{-1} + 1), \tag{4.27}$$

où $0 < B_0$, $B_1 < +\infty$, quand $0 < \epsilon' \leq \overline{\epsilon}$ et $0 < \eta' \leq \overline{\eta}$, B_0 et B_1 étant indépendant de ν.

Preuve.

– Nous montrons tout d'abord que $\{u^l\}$ est infinie. En effet, autrement il existe les nombres $\bar{l}$ et $\bar{k}$ tels que $k \in K_{\bar{l}}$ quand $k \geq \bar{k}$. Maintenant, en utilisant (4.5) et (4.17) nous avons

$$\sigma_k \|g^k\| = \frac{\langle g^k, x^k - y^k \rangle}{\|g^k\|} > \frac{\theta \epsilon_{\bar{l}} \eta_{\bar{l}}}{C'} \equiv \widetilde{\sigma} > 0$$

et le résultat suit du Lemme 4.3.3 (4).

– Ensuite, par (4.10), (4.17) et (4.26), le nombre de pas intérieurs pour n'importe quels k et l fixés n'excède pas la valeur

$$C = (C'/((1-\theta)\eta'))^2 \geq |I_{k,l}| - 1. \tag{4.28}$$

– Nous procédons maintenant à montrer que

$$\|u^l - x^*\| \leq B_0 \nu^l, \quad l = 0, 1, \ldots, \quad 0 \leq B_0 < +\infty. \tag{4.29}$$

Par le Lemme 4.3.3 (2), $\{x^k\}$ est bornée, d'où

$$\|\overline{u} - u^l\| \leq B_3 < +\infty, \quad l = 1, 2, \ldots$$

Par conséquent, par (4.17) et (4.19) nous avons

$$\begin{aligned}
\eta_l B_3 + \epsilon_l C' &\geq \eta_l B_3 + \epsilon_l D_l \\
&\geq \eta_l \|u^l - \overline{u}\| + \epsilon_l D_l \\
&\geq \tau_l(\langle \overline{g}, u^l - \overline{u} \rangle - \epsilon_l \|\overline{g}\|) - (1 - \tau_l) h(\overline{u}) \\
&\geq \tau_l(-(\epsilon_l + B_3)\|\overline{g}\|) - h(\overline{u}) + \tau_l h(\overline{u})
\end{aligned}$$

i.e.

$$\tau_l(-h(\overline{u}) + (\epsilon_l + B_3)\|\overline{g}\|) \geq -h(\overline{u}) - \eta_l B_3 - \epsilon_l C'.$$

En posant

$$\eta_l \leq \widetilde{\eta} = -\frac{h(\overline{u})}{4B_3}, \quad \epsilon' \leq \overline{\epsilon} = -\frac{h(\overline{u})}{4C'}$$

nous avons $\epsilon_l \leq \overline{\epsilon}$ et

$$\begin{aligned}
\tau_l &\geq (-h(\overline{u}) - \eta_l B_3 - \epsilon_l C')/(-h(\overline{u}) + (\epsilon_l + B_3)\|\overline{g}\|) \\
&\geq (-h(\overline{u}) + h(\overline{u})/4 + h(\overline{u})/4)/(-h(\overline{u}) + (\overline{\epsilon} + B_3)\|\overline{g}\|) \\
&\geq -h(\overline{u})/(2(-h(\overline{u}) + (\overline{\epsilon} + B_3)\|\overline{g}\|)) \\
&\equiv \tau' \geq 0
\end{aligned}$$

pour $l = 1, 2, \ldots$. En utilisant (4.18), le fait que $\tau_l \in [0, 1]$ et (4.17), nous avons

$$\begin{aligned}
(\tau_l \mu - \eta_l)\|x^* - u^l\| &\leq \epsilon_l(\tau_l \mu + D_l) \\
&\leq \epsilon_l(\mu + C').
\end{aligned}$$

Par conséquent, en posant

$$\eta_l = \eta' \leq \overline{\eta} = \min\{\widetilde{\eta}, \frac{\tau'\mu}{2}\},$$

nous voyons que (4.29) a lieu avec

$$B_0 = \max\left\{\|u^0 - x^*\|, \frac{\overline{\epsilon}(\mu + C')}{\tau'\mu - \overline{\eta}}\right\}.$$

– Ensuite, soit k quelconque tel que $k(l) < k < k(l+1)$ et $d = k - k(l)$. Nous avons alors que $\sigma_k > 0$ et en utilisant (4.5), (4.17) et (4.26)

$$\sigma_k\|g^k\| \geq \frac{\theta\epsilon_{l+1}\eta_{l+1}}{\|g^k\|} \geq \frac{\theta\epsilon_{l+1}\eta'}{C'}.$$

En combinant ces inégalités avec (4.12), (4.15) et (4.26), nous avons

$$\begin{aligned} 0 &\leq \|x^{k+1} - x^*\|^2 \\ &\leq \|x^{k(l)} - x^*\|^2 - \gamma(2-\gamma)\sum_{s=k(l)}^{k}(\sigma_s\|g^s\|)^2 \\ &\leq \|u^l - x^*\|^2 - d\gamma(2-\gamma)(\theta\epsilon_{l+1}\eta'/C')^2. \end{aligned}$$

En appliquant (4.29) dans cette inégalité, nous avons

$$\begin{aligned} |K_{l+1}| &\leq \frac{\|u^l - x^*\|^2(C')^2}{\gamma(2-\gamma)(\theta\epsilon_{l+1}\eta')^2} \\ &\leq \frac{(B_0C')^2\nu^{2l}}{\gamma(2-\gamma)(\theta\epsilon'\nu^{l+1}\eta')^2} \\ &\leq \frac{((B_0C')/(\theta\epsilon'\eta'))^2}{\nu^2\gamma(2-\gamma)} \equiv B_1'\nu^{-2}. \end{aligned} \tag{4.30}$$

D'un autre côté, de (4.29), nous avons

$$l(\delta) \leq ln(B_0/\delta)/ln(\nu^{-1}). \tag{4.31}$$

En appliquant cette inégalité avec (4.28) et (4.30) à (4.25) nous avons

$$N(\delta) \leq CB_1'\nu^{-2}(ln(B_0/\delta)/ln(\nu^{-1}) + 1)$$

i.e. la relation (4.27) a lieu avec $B_1 = CB_1'$. □

Remarque
L'assertion du Théorème 4.4.1 reste valide sans l'hypothèse additionnelle de monotonicité de F si $C = \mathbb{R}^n$. En fait, en tenant compte de (4.20) et (4.23) nous avons alors $\tau_l = 1$

et (4.29) suit directement de (4.18) et (4.26).

Donc, notre méthode atteint une estimation de complexité logarithmique, qui correspond à un taux linéaire de convergence par rapport aux pas intérieurs. Nous établissons à présent une borne supérieure similaire pour $N(\delta)$ dans le cas univoque.

Théorème 4.4.2 *Supposons que $C = I\!R^n$ et que les hypothèses (HA) soient satisfaites avec $\epsilon = 2$. Supposons que F soit Lipschitz continu de constante L. Soit $\{x^k\}$ une suite générée par la Méthode 4.1.1 où*

$$\epsilon_l = \nu^l \epsilon', \eta_l = \nu^l \eta', l = 0, 1, \ldots ; \quad \epsilon' > 0, \eta' > 0; \quad \nu \in (0,1) \tag{4.32}$$

Alors,

$$N(\delta) \leq B_1 \nu^{-6} (\ln(B_0/\delta)/\ln \nu^{-1} + 1), \tag{4.33}$$

où $0 < B_0$, $B_1 < +\infty$, B_0 et B_1 étant indépendants de ν.

Preuve.
En utilisant le même argument que celui dans la preuve du Théorème 4.3.1, nous déduisons que $\{u^l\}$ est infinie. Nous procédons maintenant à obtenir les relations similaires à (4.28)-(4.31). Nous considérons la k-ième itération de la Méthode 4.1.1, $k(l) < k < k(l+1)$. Dû aux hypothèses du Théorème, $C^* = \{x^*\}$. Alors nous devons avoir

$$\|x^k - x^*\| \leq \|u^l - x^*\| \leq \|F(u^l)\|/\mu \tag{4.34}$$

et

$$\mu \|x^k - x^*\| \leq \|F(x^k)\| \leq L\|x^k - x^*\| \tag{4.35}$$

dû a (4.12), (HA.1), (HA.2) et le fait que $F(x^*) = 0$. D'un autre côté, puisque

$$\|x^k - w^i\| \leq \epsilon_{l+1}$$

pour $i \in I_{k,l+1}$, nous voyons que

$$\begin{aligned} \|q^i - F(x^k)\| &= \|F(w^i) - F(x^k)\| \\ &\leq L\|w^i - x^k\| \\ &\leq L\epsilon_{l+1} \end{aligned} \tag{4.36}$$

et, en tenant compte de (4.20),

$$\|p^s - F(x^k)\| \leq L\epsilon_{l+1}, \quad \text{pour } s \in I_{k,l+1}. \tag{4.37}$$

Il suit de (4.11) et (4.34)-(4.36) que

$$\begin{aligned} D_{k,l} &\leq \|F(x^k)\| + L\epsilon_{l+1} \\ &\leq L(\|u^l - x^*\| + \epsilon_{l+1}) \\ &\leq L(\|F(u^l)\|/\mu + \epsilon_{l+1}). \end{aligned} \tag{4.38}$$

D'un autre côté, à la $k(l)$-ième itération, il existe un nombre $\overline{s}$ tel que $\|p^{\overline{s}}\| \leq \eta_l$ (cfr (4.4)) et $\|p^{\overline{s}} - F(u^l)\| \leq L\epsilon_l$ car $\|w^i - u^l\| \leq \epsilon_l$ et F est Lipschitz continu. D'où

$$\|F(u^l)\| \leq \|p^{\overline{s}}\| + L\epsilon_l \leq \eta_l + L\epsilon_l. \tag{4.39}$$

En utilisant cette inégalité dans (4.38) nous avons

$$\begin{aligned} D_{k,l} &\leq L((\eta_l + L\epsilon_l)/\mu + \epsilon_{l+1}) \\ &\leq L((\eta'\nu^l + L\nu^l\epsilon')/\mu + \nu^{l+1}\epsilon') \\ &\leq L((\eta' + L\epsilon')/(\mu\nu) + \epsilon')\nu^{l+1}. \end{aligned}$$

En combinant cette inégalité avec (4.10), (4.32) et en prenant en compte le critère d'arrêt (4.4), nous obtenons

$$|I_{k,l}| - 1 \leq C''\nu^{-2} \tag{4.40}$$

où

$$C'' = (L((\eta' + L\epsilon')/\mu + \epsilon')/((1-\theta)\eta'))^2.$$

Ensuite, par (4.34), (4.33) et (4.39), nous obtenons

$$\|u^l - x^*\| \leq (\eta' + L\epsilon')\nu^l/\mu,$$

i.e., (4.29) a lieu avec

$$B_0 = \max\{\|u^0 - x^*\|, (\eta' + L\epsilon')/\mu\}.$$

En utilisant (4.5), (4.29), (4.32), et (4.34)-(4.38) nous avons

$$\begin{aligned} \sigma_k\|g^k\| &= \langle g^k, x^k - y^k\rangle/\|g^k\| \\ &\geq \theta\epsilon_{l+1}\eta_{l+1}/\|g^k\| \\ &\geq \theta\epsilon_{l+1}\eta_{l+1}/(\|F(x^k)\| + L\epsilon_{l+1}) \\ &\geq \theta\epsilon_{l+1}\eta_{l+1}/(L(\|u^l - x^*\| + \epsilon_{l+1})) \\ &\geq \theta\epsilon'\eta'\nu^{l+2}/(L(B_0 + \epsilon')) \\ &\equiv B_4'\nu^{l+2}. \end{aligned}$$

En appliquant cette inégalité dans (4.12) et en prenant en compte (4.15) et (4.32), nous avons

$$\begin{aligned} 0 &\leq \|x^{k+1} - x^*\|^2 \\ &\leq \|x^{k(l)} - x^*\| - \gamma(2-\gamma)\sum_{s=k(l)}^{k}(\sigma_s\|g^s\|)^2 \\ &\leq \|u^l - x^*\|^2 - d\gamma(2-\gamma)(B_4'\nu^{l+2})^2 \end{aligned}$$

pour $k = k(l) + d$. En tenant compte de (4.29) nous avons maintenant (4.31) et

$$|K_{l+1}| \leq \nu^{-4}(B_0/B_4')^2/(\gamma(2-\gamma)) \equiv B_4\nu^{-4}.$$

En appliquant cette inégalité avec (4.31) et (4.40) nous avons

$$N(\delta) \le C'' B_4 \nu^{-6} (\ln(B_0/\delta)/\ln \nu^{-1} + 1),$$

i.e., la relation (4.33) a lieu avec $B_1 = C'' B_4$. La preuve est complète. $\square$

4.5 Modifications

La Méthode 4.1.1 peut être ajustée pour résoudre $(MVIP)$ dans le cas général où l'ensemble admissible C n'est pas associé à une certaine fonction h. En effet, nous pouvons considérer les hypothèses (HM) suivantes :

1. C est un sous-ensemble convexe, fermé et non vide de $\mathbb{R}^n$ tel que $intC \neq \emptyset$;
2. $F : C \to \Pi(\mathbb{R}^n)$ est un K-mapping ;
3. $C^d \neq \emptyset$.

Nous notons $S(x,\epsilon)$ la sphère euclidienne de rayon ϵ autour de x

$$S(x,\epsilon) = \{y \in \mathbb{R}^n | \|y - x\| = \epsilon\}$$

et $B(x,\epsilon)$ la boucle euclidienne fermée de rayon ϵ autour de x

$$B(x,\epsilon) = \{y \in \mathbb{R}^n | \|y - x\| \leq \epsilon\}.$$

Nous fixons $\rho > 0$ et nous posons

$$\begin{aligned} T_\rho(x) &= N(C,x) \bigcap S(0,\rho), \\ D(x) &= convT_\rho(x) \end{aligned} \tag{4.41}$$

pour chaque $x \in \mathbb{R}^n$. Nous pouvons maintenant remplacer les opérateurs Q et P par les opérateurs suivants $\widetilde{Q} : \mathbb{R}^n \to \Pi(\mathbb{R}^n)$ et $\widetilde{P} :\to \Pi(\mathbb{R}^n)$, respectivement, qui sont définis par

$$\widetilde{Q}(x) = \begin{cases} F(x) & \text{si} \quad x \in C, \\ D(x) & \text{si} \quad x \notin C; \end{cases} \tag{4.42}$$

et

$$\widetilde{P}(x) = \begin{cases} F(x) & \text{si} \quad x \in intC, \\ conv\{F(x) \bigcup D(x)\} & \text{si} \quad x \in C \setminus intC, \\ D(x) & \text{si} \quad x \notin C. \end{cases} \tag{4.43}$$

Nous obtenons une proposition analogue à la Proposition 4.2.1.

Proposition 4.5.1 *Les opérateurs D, $\widetilde{Q}$ et $\widetilde{P}$ définis par (4.41)-(4.43) sont des K-mappings.*

Nous notons par $\widetilde{S}^*$ l'ensemble des points stationnaires de $\widetilde{P}$, i.e., $x^* \in \widetilde{S}^*$ si et seulement si $0 \in \widetilde{P}(x^*)$.

Théorème 4.5.1 *Nous avons*
$$C^* = \widetilde{S}^*.$$

Preuve.
Puisque $intC \neq \emptyset$, alors il existe un point $\overline{u} \in C$ et un nombre $\epsilon > 0$ tel que $B(\overline{u}, \epsilon) \subseteq C$ dû au caractère fermé de C. Il suit que, pour chaque $q \in T_\rho(x^*)$ et pour chaque $x^* \notin intC$,
$$\langle q, x - x^* \rangle \leq 0 \quad \forall x \in B(\overline{u}, \epsilon),$$
d'où nous avons
$$\langle q, \overline{u} - x^* \rangle \leq -\epsilon\rho.$$
Par conséquent,
$$\|q\| \geq \epsilon\rho / \|x^* - \overline{u}\| > 0 \quad \forall q \in D(x^*).$$
Donc, $0 \notin D(x^*)$ quand $x^* \notin intC$. Il suit que $\widetilde{S}^* \subseteq C$. De plus, si $x^* \in C \setminus intC$ et $0 \in \widetilde{P}(x^*)$, alors $0 \notin D(x^*)$ et $N(C, x^*) = coneD(x^*)$. Par le Lemme 4.2.1 avec $W = F(x^*)$ et $Y = D(x^*)$, la relation $0 \in \widetilde{P}(x^*)$ est alors équivalente à
$$0 \in F(x^*) + N(C, x^*).$$
Ceci est aussi vrai dans le cas où $x^* \in intC$, puisque nous avons alors $N(C, x^*) = \{0\}$. La preuve est complète. $\square$

Donc, nous pouvons remplacer Q par $\widetilde{Q}$ dans la Méthode 4.1.1 et P par $\widetilde{P}$ dans les preuves correspondantes. Alors, il est facile de voir que tous les résultats des Théorèmes 4.3.1 et 4.4.1 restent valides après de petites modifications.

Ensuite, la simple règle (4.6) dans la Méthode 4.1.1 peut être remplacée par une des suivantes :
$$p^{i+1} = Nrconv\{q^0, \dots, q^{i+1}\}, \tag{4.44}$$
ou
$$p^{i+1} = Nrconv\{p^i, q^{i+1}, S_i\}, \tag{4.45}$$
où $S_i \subseteq conv\{q^0, \dots, q^i\}$. De la définition il suit que l'assertion du Lemme 4.3.1 reste valide après cette modification et que les preuves des autres résultats ne sont pas changées. Donc, nous pouvons utiliser la règle (4.44) ou (4.45) pour accélérer la convergence de la méthode, mais nous remarquons qu'elles ont besoin de plus de mémoire et de travail par pas intérieur.

Lemme 5.1.1 *1. $S^*_{(d)}$ est convexe et fermé.*

2. $S^*_{(d)} \subseteq S^*$.

3. Si P est pseudomonotone, alors $S^*_{(d)} = S^*$.

En plus du Théorème 4.2.1, il est facile d'obtenir le résultat équivalent pour les problèmes $(DMVIP)$ et (5.1).

Proposition 5.1.1 $C^d = S^*_{(d)}$.

Preuve.

– Soient $x^* \in C^d$ et $x \in \mathbb{R}^n$. Si $h(x) \leq 0$ alors pour chaque $t \in F(x)$, nous avons $\langle t, x - x^* \rangle \geq 0$ à cause de $(DMVIP)$ et (4.7). Si $h(x) \geq 0$, alors, par définition,

$$0 \geq h(x^*) - h(x) \geq \langle t, x^* - x \rangle \quad \forall t \in \partial h(x).$$

En tenant compte de (4.7), nous obtenons $x^* \in S^*_{(d)}$. D'où, $C^d \subseteq S^*_{(d)}$.

– Inversément, par le Lemme 5.1.1 (2) et le Théorème 4.2.1, nous avons $S^*_{(d)} \subseteq S^* = C^* \subseteq C$. Supposons par contradiction qu'il existe un élément $z^* \in S^*_{(d)} \setminus C^d$. Alors, comme $z^* \notin C^d$, il existe un point $x \in C$ et un élément $t \in F(x)$ tel que $\langle t, x - z^* \rangle < 0$. Mais $F(x) \subseteq P(x)$, d'où $z^* \notin S^*_{(d)}$, une contradiction. La preuve est complète. □

Donc, sous les hypothèses (HG), P est un K-mapping et $S^*_{(d)}$ est non vide. Puisque l'inclusion multivoque (4.8) est un cas particulier de $(MVIP)$, nous pouvons appliquer la Méthode 4.1.1 au problème (4.8). De telle manière, nous obtenons également une solution du problème initial $(MVIP)$. Nous considérons de plus une version inexacte de la Méthode 4.1.1, qui admet des calculs inexacts d'éléments de $P(x^k)$. Une telle technique nous permet de prendre en compte divers calculs d'erreurs et en partie "rendre différentiable" le problème initial puisque tout élément de $P(x^k)$ peut maintenant être remplacé par l'enveloppe convexe d'éléments calculés proche de x^k. Pour chaque point $x \in \mathbb{R}^n$, poser

$$H_\delta(x) = conv \bigcup_{x \in B(\nu,\delta)} P(x).$$

Méthode 5.1.1 *Choisir un point $x^0 \in I\!R^n$, une suite $\{\gamma_j\}$ satisfaisant (4.3), une suite bornée positive $\{\delta_k\}$ et deux suites positives croissantes $\{\epsilon_l\}$ et $\{\eta_l\}$. Choisir également un nombre $\theta \in (0,1)$. Poser $k=0$, $l=1$, $u^0 = x^0$, $j(0)=0$ et $k(0)=0$.*

Pas 1 (Procédure auxiliaire $\widetilde{D}'_k$)

Pas 1.1 : Choisir q^0 dans $H_{\delta_k}(x^k)$, poser $i=0$, $p^i = q^i$ et $w^0 = x^k$.
Pas 1.2 : Si

$$\|p^i\| \leq \eta_l,$$

poser $k(l)=k$, $u^l = x^{k+1} = x^k$, $j(k+1)=j(k)$, $k=k+1$ et $l=l+1$ et retour au Pas 1.

Pas 1.3 : Poser $w^{i+1} = x^k - \epsilon_l p^i/\|p^i\|$, choisir $q^{i+1} \in H_{\delta_k}(w^{i+1})$. Si

$$\langle q^{i+1}, p^i \rangle > \theta \|p^i\|^2$$

alors poser $y^k = w^{i+1}$, $g^k = q^{i+1}$ et $j(k+1) = j(k)+1$, et aller au Pas 2.

Pas 1.4 : Poser

$$p^{i+1} = Nrconv\{p^i, q^{i+1}\}$$

$i=i+1$ et retour au Pas 1.2.

Pas 2 (Itération principale) *Poser $\sigma_k = \langle g^k, x^k - y^k \rangle / \|g^k\|^2$,*

$$x^{k+1} = x^k - \gamma_{j(k)} \sigma_k g^k \tag{5.2}$$

$k=k+1$ et retour au Pas 1.

Donc, la Méthode 5.1.1, contrairement à la Méthode 4.1.1, n'a pas besoin d'utiliser les opérateurs quasi non expansifs admissibles par rapport à C. Néanmoins, elle nous permet également d'approximer une solution de $(MVIP)$. Dans ce qui suit, nous appellerons un changement de l'indice i un pas intérieur.

2. La relation (5.3) suit du Lemme C.0.1. Par construction, $q^i \in H_{\delta_k}(w^i)$ et $w^i \in B(x^k, \epsilon_l)$, d'où nous avons $q^i \in H_{\epsilon_l+\delta_k}$. Puisque $H_{\epsilon_l+\delta_k}$ est convexe, nous obtenons alors $p^i \in H_{\epsilon_l+\delta_k}$ à cause de (5.3). Donc (5.4) est vrai.

3. Dans le cas $k = k(l)$, (5.5) a évidemment lieu. Considérons le cas $k \neq k(l)$. Fixons $x^* \in S^*_{(d)}$. Par définition, il existe les vecteurs $z^{i,j} \in B(w^i, \delta_k)$, $q^{i,j} \in P(z^{i,j})$ et les nombres μ_j, $j \in J$, tels que

$$q^i = \sum_{j \in J} \mu_j q^{i,j}, \quad \sum_{j \in J} \mu_j = 1, \quad \mu_j \geq 0, \quad j \in J.$$

D'un autre côté, en tenant compte de (5.1), nous avons $\langle q^{i,j}, z^{i,j} - x^* \rangle \geq 0$. En combinant ces relations nous avons

$$\begin{aligned} \langle q^i, w^i - x^* \rangle &= \sum_{j \in J} \mu_j (\langle q^i, z^{i,j} - x^* \rangle + \langle q^{i,j}, w^i - z^{i,j} \rangle) \\ &\geq -\sum_{j \in J} \mu_j \| q^{i,j} \| \delta_k \\ &\geq -\delta_k C_k. \end{aligned}$$

Il suit que

$$\begin{aligned} \langle g^k, x^k - x^* \rangle &= \langle g^k, x^k - y^k \rangle + \langle g^k, y^k - x^* \rangle \\ &= \sigma_k \| g^k \|^2 + \langle q^i, w^i - x^* \rangle \\ &\geq \sigma_k \| g^k \|^2 - \delta_k C_k. \end{aligned}$$

En combinant cette inégalité avec (5.2) nous avons

$$\begin{aligned} \| x^{k+1} - x^* \| &= \| x^k - \gamma_{j(k)} \sigma_k g^k - x^* \|^2 \\ &= \| x^k - x^* \|^2 - 2\gamma_{j(k)} \sigma_k \langle g^k, x^k - x^* \rangle + (\gamma_{j(k)} \sigma_k \| g^k \|)^2 \\ &\leq \| x^k - x^* \|^2 - 2\gamma_{j(k)} \sigma_k (\sigma_k \| g^k \|^2 - \delta_k C_k) + (\gamma_{j(k)} \sigma_k \| g^k \|)^2 \end{aligned}$$

i.e., (5.5) a lieu. □

Nous obtenons à présent les propriétés clés de la Méthode 5.1.1.

Proposition 5.2.1 *1. Le nombre de pas intérieurs à chaque itération de la Méthode 5.1.1 est fini.*

2. Soient les paramètres de la Méthode 5.1.1 choisis par (4.15), (4.13) et

$$\{\delta_k\} \searrow 0. \tag{5.6}$$

Si, de plus,

$$C_k \leq C' < +\infty \quad pour\ k = 0, 1, \ldots, \tag{5.7}$$

alors il existe les suites $\{x^{k_s}\}$, $\{p^{k_s}\}$ et $\{\beta_{k_s}\} \searrow 0$ telles que

$$\lim_{s \to +\infty} p^{k_s} = 0$$

et

$$p^{k_s} \in H_{\beta_{k_s}}(x^{k_s}).$$

Il est clair que (5.7) a lieu si, par exemple,

$$\sup\{\|g\| | g \in P(x), x \in \mathbb{R}^n\} \leq C' < +\infty.$$

Cependant, pour garantir la convergence, nous avons besoin d'hypothèses supplémentaires.

Corollaire 5.2.1 *Soient les paramètres de la Méthode 5.1.1 choisis par (4.13), (4.15) et (5.6). Si $\{x^k\}$ est bornée, il existe une valeur d'adhérence de $\{x^k\}$ qui appartient à S^*.*

Afin d'obtenir les résultats de convergence sous des conditions plus faibles, la Méthode 5.1.1 doit être légèrement modifiée. En fait, il faut choisir un ensemble compact convexe V tel que $V \bigcap S^*_{(d)} \neq \emptyset$ et remplacer la règle (5.2) par la suivante :

$$x^{k+1} = \pi_V(x^k - \gamma_{j(k)}\sigma_k g^k). \tag{5.8}$$

Des propriétés de projection non expansive (cfr. (A.0.3 (3)), il suit que toutes les assertions du Lemme 5.2.1 et de la Proposition 5.2.1 avec $S^*_{(d)}$ étant remplacé par $S^*_{(d)} \bigcap V$ restent valides et que $\{x^k\}$ est maintenant bornée. En combinant ces assertions avec le Corollaire 5.2.1, nous obtenons immédiatement le résultat de convergence suivant.

Chapitre 6

Méthode faisceau pour le problème (EP)

6.1 Algorithme général

Dans cette section, nous considérons le problème d'équilibre général (EP) et l'algorithme introduit par Mastroeni pour le résoudre où le paramètre $\epsilon = \epsilon_k > 0$ est autorisé à varier à chaque itération. L'algorithme peut être exprimé comme suit :

Soit $x^k \in C$ donné, trouver $x^{k+1} \in C$ la solution du problème (P_k).

Nous imposons que le gradient ∇h soit Lipschitz continu sur C de constante $L > 0$. Nous notons par $\kappa > 0$ le module de la fonction fortement convexe h.

Nous supposons, par la suite, qu'il existe au moins une solution du problème (EP).

La fonction $f(x^k, \cdot)$, notée f_k par la suite, est remplacée dans le problème (P_k) par une autre fonction convexe $\overline{f}_k$ de telle manière que le nouveau problème

$$(\overline{P}_k) \quad \min_{y \in C}\{\epsilon \overline{f}_k(y) + h(y) - h(x^k) - \langle \nabla h(x^k), y \rangle\}$$

soit plus facile à résoudre et que l'algorithme correspondant génère une suite $\{x^k\}$ qui converge vers une certaine solution du problème (EP) :

Soit $x^k \in C$ donné, trouver $x^{k+1} \in C$ la solution du problème $(\overline{P}_k)$.

Théorème 6.1.1 *Supposons que $\epsilon_k \geq \epsilon > 0$ pour tout $k \in I\!N$. Si la suite $\{x^k\}$ générée par l'algorithme général est bornée et est telle que $\|x^{k+1} - x^k\| \to 0$, $k \in I\!N$, alors chaque valeur d'adhérence de $\{x^k\}_{k \in I\!N}$ est une solution du problème (EP).*

Preuve.
Soit x^* une valeur d'adhérence de $\{x^k\}_{k\in\mathbb{N}}$ et soit $\{x^k\}_{k\in\mathbb{K}\subset\mathbb{N}}$ une certaine sous-suite convergeant vers x^*. Puisque $\|x^{k+1} - x^k\| \to 0$, nous avons aussi $\{x^{k+1}\}_{k\in\mathbb{K}} \to x^*$. D'où, comme $\overline{f}_k \leq f_k$ et $f_k(x^{k+1}) \leq \mu \overline{f}_k(x^{k+1})$, nous obtenons

$$\frac{1}{\mu} f_k(x^{k+1}) \leq \overline{f}_k(x^{k+1}) \leq f_k(x^{k+1}).$$

Maintenant $f_k(x^{k+1}) = f(x^k, x^{k+1}) \to f(x^*, x^*) = 0$ pour $k \to +\infty$ parce que $x^k \to x^*$, $x^{k+1} \to x^*$ pour $k \to +\infty$, $k \in \mathbb{K}$, et f est continue. D'où $\overline{f}_k(x^{k+1}) \to 0$ pour $k \to +\infty$. D'un autre côté, puisque x^{k+1} résout le problème d'optimisation convexe $(\overline{P}_k)$, nous avons

$$0 \in \partial\{\epsilon_k(\overline{f}_k) + \psi_C\}(x^{k+1}) - \nabla h(x^k) + \nabla h(x^{k+1})$$

i.e.,

$$\nabla h(x^{k+1}) - \nabla h(x^k) \in \partial\{\epsilon_k(\overline{f}_k) + \psi_C\}(x^{k+1})$$

où ψ_C est la fonction indicatrice associée à C (i.e. $\psi_C(x) = 0$ si $x \in C$ et $+\infty$ sinon). En utilisant la définition de sous-différentielle, nous obtenons

$$\forall y \in C \quad \overline{f}_k(y) - \overline{f}_k(x^{k+1}) \geq \frac{1}{\epsilon_k}\langle \nabla h(x^k) - \nabla h(x^{k+1}), y - x^{k+1}\rangle. \tag{6.2}$$

En appliquant l'inégalité de Cauchy-Schwarz et les propriétés $\overline{f}_k \leq f_k$ et ∇h est Lipschitz continu sur C de constante $L > 0$, nous obtenons successivement pour tout $y \in C$,

$$\begin{aligned} f_k(y) - \overline{f}_k(x^{k+1}) &\geq -\frac{1}{\epsilon_k}\|\nabla h(x^k) - \nabla h(x^{k+1})\|\|y - x^{k+1}\| \\ &\geq -\frac{L}{\epsilon_k}\|x^k - x^{k+1}\|\|y - x^{k+1}\|. \end{aligned}$$

En prenant la limite sur $k \in K$, nous déduisons

$$\forall y \in C \quad f(x^*, y) \geq 0$$

parce que f est continue, $\overline{f}_k(x^{k+1}) \to 0$, $\|x^k - x^{k+1}\| \to 0$, $\|y - x^{k+1}\| \to \|y - x^*\|$ et $\epsilon_k \geq \epsilon > 0$. Mais cela signifie que x^* est une solution du problème (EP). □

Dans le théorème suivant, nous donnons les conditions pour obtenir que la suite $\{x^k\}$ soit bornée et que $\|x^{k+1} - x^k\| \to 0$.

Théorème 6.1.2 *Supposons qu'il existe* $\gamma, c, d > 0$ *et une fonction positive* $g : C \times C \to I\!R$ *telle que pour tout* $x, y, z \in C$,

1. $f(x,y) \geq 0 \Rightarrow f(y,z) \leq -\gamma g(x,y)$;
2. $f(x,z) - f(y,z) - f(x,y) \leq cg(x,y) + d\|z-y\|^2$.

Si la suite $\{\epsilon_k\}_{k\in I\!N}$ *est décroissante et* $\epsilon_k < \frac{\beta\mu}{2d}$ *pour tout* k *et si* $\frac{c}{\gamma} \leq \mu \leq 1$, *alors la suite* $\{x^k\}_{k\in I\!N}$ *générée par l'algorithme général est bornée et* $\lim_{k\to+\infty} \|x^{k+1} - x^k\| = 0$.

Preuve.
Soit x^* une solution du problème (EP) et considérons pour chaque $k \in \mathbb{N}$ la fonction de Lyapounov $\Gamma^k : C \times C \to \mathbb{R}$ définie pour tout $y, z \in C$, par

$$\Gamma^k(y,z) = h(z) - h(y) - \langle \nabla h(y), z - y\rangle + \frac{\epsilon_k}{\mu} f(z,y). \tag{6.3}$$

Puisque h est fortement convexe de module $\kappa > 0$, nous avons immédiatement que, pour tout $x \in C$,

$$\Gamma^k(x^k, x^*) \geq \frac{\kappa}{2}\|x^k - x^*\|^2. \tag{6.4}$$

Remarquant que $\epsilon_{k+1} \leq \epsilon_k$ pour tout $k \in \mathbb{N}$, la différence $\Gamma^{k+1}(x^{k+1}, x^*) - \Gamma^k(x^k, x^*)$ peut alors être évaluée comme suit :

$$\begin{aligned}
\Gamma^{k+1}(x^{k+1}, x^*) - \Gamma^k(x^k, x^*) &\leq h(x^k) - h(x^{k+1}) + \langle \nabla h(x^k), x^{k+1} - x^*\rangle \\
&\quad + \langle \nabla h(x^k) - \nabla h(x^{k+1}), x^* - x^{k+1}\rangle \\
&\quad + \frac{\epsilon_k}{\mu}\{f(x^*, x^{k+1}) - f(x^*, x^k)\} \\
&= s_1 + s_2 + s_3
\end{aligned} \tag{6.5}$$

avec

$$\begin{aligned}
s_1 &= h(x^k) - h(x^{k+1}) + \langle \nabla h(x^k), x^{k+1} - x^*\rangle \\
s_2 &= \langle \nabla h(x^k) - \nabla h(x^{k+1}), x^* - x^{k+1}\rangle \\
s_3 &= \frac{\epsilon_k}{\mu}\{f(x^*, x^{k+1}) - f(x^*, x^k)\}.
\end{aligned}$$

Pour s_1, nous obtenons facilement de la forte convexité de h que

$$s_1 \leq -\frac{\kappa}{2}\|x^{k+1} - x^k\|^2. \tag{6.6}$$

Pour s_2, nous obtenons, en prenant $y = x^*$ dans (6.2)

$$\begin{aligned}
s_2 &= \langle \nabla h(x^k) - \nabla h(x^{k+1}), x^* - x^{k+1}\rangle \\
&\leq \epsilon_k\{\overline{f}_k(x^*) - \overline{f}_k(x^{k+1})\} \\
&\leq \epsilon_k\left\{f(x^k, x^*) - \frac{1}{\mu} f(x^k, x^{k+1})\right\}
\end{aligned}$$

parce que $\overline{f}_k \leq f(x^k, \cdot)$ et (6.1) a lieu. Alors, en utilisant l'hypothèse (2), nous déduisons que

$$\begin{aligned} s_2 + s_3 &\leq \epsilon_k \left\{ f(x^k, x^*) - \frac{1}{\mu} f(x^k, x^{k+1}) \right\} + \frac{\epsilon_k}{\mu} \{ f(x^*, x^{k+1}) - f(x^*, x^k) \} \\ &= \frac{\epsilon_k}{\mu} \{ f(x^*, x^{k+1}) - f(x^*, x^k) - f(x^k, x^{k+1}) \} + \epsilon_k f(x^k, x^*) \\ &\leq \frac{\epsilon_k}{\mu} \{ c g(x^*, x^k) + d \|x^{k+1} - x^k\|^2 \} + \epsilon_k f(x^k, x^*). \end{aligned}$$

Par conséquent, nous avons

$$\begin{aligned} \Gamma^{k+1}(x^{k+1}, x^*) - \Gamma^k(x^k, x^*) &\leq \epsilon_k f(x^k, x^*) - \frac{1}{2} \left(\kappa - 2\frac{\epsilon_k d}{\mu} \right) \|x^{k+1} - x^k\|^2 \\ &\quad + \frac{\epsilon_k c}{\mu} g(x^*, x^k). \end{aligned}$$

En appliquant l'hypothèse (1) avec $x = x^*$ et $y = x^k$, puisque $f(x^*, x^k) \geq 0$, nous obtenons

$$f(x^k, x^*) \leq -\gamma(x^*, x^k).$$

Enfin, nous avons que

$$\begin{aligned} \Gamma^{k+1}(x^{k+1}, x^*) - \Gamma^k(x^k, x^*) &\leq -\frac{1}{2} \left(\kappa - 2\frac{\epsilon_k d}{\mu} \right) \|x^{k+1} - x^k\|^2 \\ &\quad - \epsilon_k \left(\gamma - \frac{c}{\mu} \right) g(x^*, x^k). \end{aligned} \tag{6.7}$$

Puisque $\epsilon_k < \frac{\kappa\mu}{2d}$ pour tout k et $\mu \geq \frac{c}{\gamma}$, par (6.4) et (6.7), il suit que $\{\Gamma^k(x^k, x^*)\}_{k \in \mathbb{N}}$ est une suite décroissante bornée inférieurement par 0. D'où, elle est convergente dans $\mathbb{R}$. En utilisant encore (6.4), nous déduisons que la suite $\{x^k\}_{k \in \mathbb{N}}$ est bornée et, en passant à la limite dans (6.7), que la suite $\{\|x^{k+1} - x^k\|\}_{k \in \mathbb{N}}$ converge vers zéro. □

En combinant les Théorèmes 6.1.1 et 6.1.2, nous déduisons le théorème suivant.

Théorème 6.1.3 *Supposons que $\epsilon_k \geq \epsilon > 0$ pour tout $k \in \mathbb{N}$ et que toutes les hypothèses du Théorème 6.1.2 soient satisfaites, alors la suite $\{x^k\}_{k \in \mathbb{N}}$ générée par l'algorithme général converge vers une solution du problème (EP).*

Donc, dans le but d'obtenir la convergence de l'algorithme général, nous avons besoin des conditions (1) et (2).

- La condition (1) est une condition de monotonicité.
  En effet, quand $g = 0$, cette condition signifie que f est pseudomonotone (i.e. $f(x,y) \geq 0 \Rightarrow f(y,x) \leq 0$) et quand $g(x,y) = \|x-y\|^2$ que f est fortement pseudomonotone de module γ (i.e. $f(x,y) \geq 0 \Rightarrow f(y,x) \leq -\gamma\|x-y\|^2$).

- La condition (2) est une condition du type Lipschitz.

- Le lien entre les conditions (1) et (2) est fait par la fonction g dont le choix dépend de la structure du problème. Donc, par exemple quand $f(x,y) = \varphi(x) - \varphi(y)$ avec $\varphi : C \to \mathbb{R}$ une fonction convexe continue, i.e., quand le problème (EP) est un problème d'optimisation convexe avec contraintes, il suffit de choisir $g(x,y) = 0$ pour tout $x, y \in C$ pour obtenir que les conditions (1) et (2) soient satisfaites.

Une autre condition suffisante pour obtenir les conditions (1) et (2) est donnée dans la proposition suivante.

Proposition 6.1.1 *Si f est fortement monotone et si (1.4) a lieu, alors les conditions (1) et (2) sont satisfaites avec* $g(x,y) = \|x-y\|^2$.

Preuve.
Si $f(x,y) \geq 0$, alors par la forte monotonicité de f, nous avons

$$f(y,x) \leq -f(x,y) - \gamma\|x-y\|^2 \leq -\gamma\|x-y\|^2 = -\gamma g(x,y).$$

La condition (2) est immédiate de (1.4). □

Comme conséquence de cette proposition, le Théorème 6.1.3 est aussi valide sous les hypothèses (1.3) et (1.4).

Nous considérons à présent le cas où f est donnée par (1.5) et nous introduisons la définition suivante :

Définition 6.1.2 *F est φ-co-coercif sur C s'il existe $\gamma > 0$ tel que pour tout $x, y \in C$, si $\langle F(x), y-x\rangle + \varphi(y) - \varphi(x) \geq 0$, alors*

$$\langle F(y), y-x\rangle + \varphi(y) - \varphi(x) \geq \gamma\|F(y) - F(x)\|^2. \tag{6.8}$$

Il est facile de prouver que si F est co-coercif sur C, alors F est φ-co-coercif sur C. En effet, si F est co-coercif sur C, alors il existe $\gamma > 0$ tel que

$$\forall x, y \in C \quad \langle F(x) - F(y), x-y\rangle \geq \gamma\|F(x) - F(y)\|^2.$$

Mais alors, si $\langle F(y), y-x\rangle + \varphi(y) - \varphi(x) \geq 0$, nous avons

$$\begin{aligned}\langle F(y), y-x\rangle + \varphi(y) - \varphi(x) &= \langle F(y)-F(x), y-x\rangle + \langle F(x), y-x\rangle \\ &\quad + \varphi(y) - \varphi(x) \\ &\geq \gamma\|F(y)-F(x)\|^2\end{aligned}$$

i.e. l'inégalité (6.8).

Dans la proposition suivante, un autre choix de la fonction g est nécessaire pour obtenir (1) et (2).

Proposition 6.1.2 *Soit* $f(x,y) = \langle F(x), y-x\rangle + \varphi(y) - \varphi(x)$ *où* $F : C \to I\!R^n$ *est continu et* $\varphi : C \to I\!R$ *est convexe. Si* F *est* φ*-co-coercif sur* C*, alors il existe une fonction positive* $g : C \times C \to I\!R$ *et* $\gamma > 0$ *tels que pour tout* $x, y, z \in C$ *et pour tout* $\nu > 0$,

$$f(x,y) \geq 0 \Rightarrow f(y,x) \leq -\gamma g(x,y),$$

$$f(x,z) - f(y,z) - f(x,y) \leq \tfrac{1}{2\nu} g(x,y) + \tfrac{\nu}{2}\|z-y\|^2.$$

Preuve.
En utilisant la définition de f et la φ-co-coercivité de F sur C, nous avons qu'il existe $\gamma > 0$ tel que pour tout $x \in C$

$$f(x,y) \leq 0 \Rightarrow f(y,x) \leq -\gamma\|F(y)-F(x)\|^2.$$

D'un autre côté, nous avons pour n'importe quel $\nu > 0$,

$$f(x,z) - f(y,z) - f(x,y) = \langle F(x)-F(y), z-y\rangle \leq \frac{1}{2\nu}\|F(x)-F(y)\|^2 + \frac{\nu}{2}\|z-y\|^2.$$

Donc, avec $g(x,y) = \|F(y)-F(x)\|^2$, nous obtenons les deux inégalités. $\square$

6.2 Algorithme faisceau

Afin d'obtenir un algorithme implémentable, nous devons maintenant dire comment construire une μ-approximation $\overline{f}_k$ de f_k en x^k tel que le problème $(\overline{P}_k)$ soit plus facile à résoudre que le problème (P_k). Nous supposons ici que $\mu \in (0,1)$. Dans ce but, nous observons que si $\overline{f}_k$ est une fonction linéaire par morceaux de la forme

$$\overline{f}_k(y) = \max_{1 \leq j \leq p} \{\langle a_j, y\rangle + b_j\},$$

où $a_j \in \mathbb{R}^n$, $b_j \in \mathbb{R}$ pour $j = 1, \ldots, p$, le problème $(\overline{P}_k)$ est équivalent au problème

$$(QP_k) = \begin{cases} \min & \{\epsilon_k v + h(y) - h(x^k) - \langle \nabla h(x^k), y - x^k\rangle\} \\ s.c & v \geq \langle a_j, y\rangle + b_j, \quad j = 1, \ldots, p \\ & y \in C. \end{cases}$$

Il existe beaucoup de méthodes numériques efficaces pour résoudre un tel problème. Quand $\overline{f}_k$ est une fonction convexe linéaire par morceaux, il est judicieux de construire $\overline{f}_k$ morceau par morceau, en générant des modèles successifs

$$\overline{f}_k^i \quad i = 1, 2, \ldots$$

jusqu'à ce que (si possible) $\overline{f}_k^{i_k}$ soit une μ-approximation en x^k pour certains $i_k \geq 1$.

Pour $i = 1, 2, \ldots$, nous notons par y_k^i l'unique solution du problème

$$(P_k^i) \quad \min_{y \in C}\{\epsilon \overline{f}(x^k, y) + h(y) - h(x^k) - \langle \nabla h(x^k), y\rangle\}$$

et nous posons $\overline{f}_k = \overline{f}_k^{i_k}$ et $x^{k+1} = y_k^{i_k}$.

Afin d'obtenir une μ-approximation $\overline{f}_k^{i_k}$ de f_k en x^k, nous devons imposer certaines conditions sur les modèles successifs $\overline{f}_k^i$, $i = 1, 2, \ldots$.

Nous définissons les fonctions affines l_k^i, $i = 1, 2, \ldots$ par

$$l_k^i(y) = \overline{f}_k^i(y_k^i) + \langle \gamma_k^i, y - y_k^i\rangle \quad \forall y \in C$$

où $\gamma_k^i = \frac{1}{\epsilon_k}[\nabla h(x^k) - \nabla h(y_k^i)]$. Par l'optimalité de y_k^i, nous avons

$$\gamma_k^i \in \partial(\overline{f}_k^i + \delta_C(x))(y_k^i). \tag{6.9}$$

Nous observons alors que

$$l_k^i(y_k^i) = \overline{f}_k^i(y_k^i) \quad \text{et} \quad l_k^i(y) \leq \overline{f}_k^i(y) \quad \forall y \in C. \tag{6.10}$$

Nous supposons à présent que les conditions suivantes soient satisfaites par les modèles convexes $\overline{f}_k^i$, (C1) $\overline{f}_k^i \leq f_k$ sur C pour $i = 1, 2, \ldots$

(C2) $\overline{f}_k^{i+1} \geq f_k(y_k^i) + \langle s(y_k^i), \cdot - y_k^i \rangle$ sur C pour $i = 1, 2, \ldots$

(C3) $\overline{f}_k^{i+1} \geq l_k^i$ sur C pour $i = 1, 2, \ldots,$

où $s(y_k^i)$ est le sous-gradient de f_k en y_k^i.

Plusieurs modèles satisfont ces conditions. Par exemple, pour le premier modèle $\overline{f}_k^i$, nous pouvons prendre la fonction linéaire

$$\overline{f}_k^i(y) = f_x(x^k) + \langle s(x^k), y - x^k \rangle \quad \text{pour tout } y \in C.$$

Puisque $s(x^k) \in \partial f_k(x^k)$, la conditon (C1) est satisfaite pour $i = 1$. Pour les modèles suivants $\overline{f}_k^i$, $i = 2, \ldots$, il existe plusieurs possibilités. Un premier exemple est de prendre pour $i = 1, 2, \ldots$

$$\overline{f}_k^{i+1}(y) = \max\{l_k^i(y), f_k(y_k^i) + \langle s(y_k^i), y - y_k^i \rangle\}. \tag{6.11}$$

Les conditions (C2) et (C3) sont évidemment satisfaites et la condition (C1) est aussi satisfaite pour $i = 2, 3, \ldots$, parce que chaque morceau linéaire de ces fonctions sont inférieurs à f_k. Un autre exemple est de prendre pour $i = 1, 2, \ldots$

$$\overline{f}_k^{i+1}(y) = \max_{0 \leq j \leq i} \{f_k(y_k^j) + \langle s(y_k^j), y - y_k^j \rangle\}. \tag{6.12}$$

où $y_k^0 = x^k$. Puisque $s(y_k^j) \in \partial f_k(y_k^j)$ pour $j = 0, \ldots, i$ et puisque $\overline{f}_k^{i+1} \geq \overline{f}_k^i \geq l_k^i$, il est facile de voir que les conditions (C1)-(C3) sont satisafites. En comparant (6.11) et (6.12) nous pouvons dire que l_k^i joue le même rôle que les i fonctions linéaires $f_k(y_k^i) + \langle s(y_k^i), y - y_k^i \rangle$, $j = 0, \ldots, i-1$. C'est la raison pour laquelle cette fonction l_k^i est appelée la fonction aggrégée affine. Le premier exemple (6.11) est intéressant d'un point de vue numérique, parce que son emploi permet de limiter le nombre de contraintes linéaires dans les sous-problèmes (QP_k).

L'algorithme permettant de passer de x^k à x^{k+1}, i.e., pour faire ce que nous appelons un pas sérieux, peut être exprimé de la manière suivante.

Algorithme 6.2.1 ***Algorithme de pas sérieux :***

Soit $x^k \in C$ et $\mu \in (0, 1)$. Poser $i = 1$.

Pas 1. *Choisir $\overline{f}_k^i$ une fonction convexe qui satisfait (C1)-(C3) et résoudre le problème (P_k^i) pour obtenir y_k^i.*

Pas 2*. Si $f_k(y_k^i) \leq \mu \overline{f}_k^i(y_k^i)$, alors poser $x^{k+1} = y_k^i$ et $i_k = i$ et STOP ; x^{k+1} est un pas sérieux.*

Pas 3*. $i = i + 1$ et retour au Pas 1.*

Notre but est maintenant de prouver que si x^k n'est pas une solution du problème (EP) et si les modèles $\overline{f}_k^i$, $i = 1, \ldots$ satisfont (C1)-(C3), alors il existe $i_k \geq 1$ tel que $\overline{f}_k^{i_k}$ est une μ-approximation de f_k en x^k, i.e., que STOP a lieu au Pas 2 après un nombre fini d'itérations.

Pour prouver cela, nous avons besoin d'un lemme dont la preuve utilise les fonctions suivantes :

$$\begin{aligned}
\widetilde{l}_k^i(y) &= l_k^i(y) + \frac{1}{\epsilon_k}\{h(y) - h(x^k) - \langle \nabla h(x^k), y - x^k \rangle\}, \\
\widetilde{f}_k^i(y) &= \overline{f}_k^i(y) + \frac{1}{\epsilon_k}\{h(y) - h(x^k) - \langle \nabla h(x^k), y - x^k \rangle\}.
\end{aligned}$$

En utilisant (6.9) et (6.10), nous obtenons :

$$\begin{aligned}
\widetilde{l}_k^i(y) - \widetilde{l}_k^i(y_k^i) &= l_k^i(y) + \frac{1}{\epsilon_k}\{h(y) - h(x^k) - \langle \nabla h(x^k), y - x^k \rangle\} \\
&\quad - l_k^i(y_k^i) - \frac{1}{\epsilon_k}\{h(y_k^i) - h(x^k) - \langle \nabla h(x^k), y_k^i - x^k \rangle\} \\
&= \overline{f}_k^i(y_k^i) + \langle \gamma_k^i, y - y_k^i \rangle + \frac{1}{\epsilon_k}\{h(y) - h(x^k) - \langle \nabla h(x^k), y - x^k \rangle\} \\
&\quad - \overline{f}_k^i(y_k^i) - \frac{1}{\epsilon_k}\{h(y_k^i) - h(x^k) - \langle \nabla h(x^k), y_k^i - x^k \rangle\} \\
&= \frac{1}{\epsilon_k}\{h(x) - h(y_k^i) - \langle \nabla h(y_k^i), y - y_k^i \rangle\}.
\end{aligned}$$

Par conséquent, nous obtenons

$$\widetilde{l}_k^i(y) = \widetilde{l}_k^i(y_k^i) + \frac{1}{\epsilon_k}\{h(x) - h(y_k^i) - \langle \nabla h(y_k^i), y - y_k^i \rangle\}. \tag{6.13}$$

De plus par (6.9) et (C3), nous avons

$$\widetilde{f}_k^i(x^k) = \overline{f}_k^i(x^k)\,; \tag{6.14}$$

$$\widetilde{l}_k^i(y_k^i) = \widetilde{f}_k^i(y_k^i)\,; \tag{6.15}$$

$$\widetilde{l}_k^i \leq \widetilde{f}_k^{i+1} \quad \text{sur } C. \tag{6.16}$$

Lemme 6.2.1 *Supposons que les modèles $\overline{f}_k^i$, $i \in I\!N_0$ satisfont (C1)-(C3) et soit, pour chaque i, y_k^i l'unique solution du problème (P_k^i). Alors*

1. $f_k(y_k^i) - \overline{f}_k^i(y_k^i) \to 0$,
2. $y_k^i \to \overline{y}_k \equiv \arg\min_{y \in C}\{\epsilon_k f_k(y) + h(y) - h(x^k)\langle \nabla h(x^k), y - x^k\rangle\}$,

où $i \to +\infty$.

Preuve

1. Pour obtenir la première assertion, nous utilisons les 3 étapes suivantes.

   (a) <u>La suite $\{\widetilde{l}_k^i(y_k^i)\}_{i\in\mathbb{N}_0}$ est convergente et $y_k^{i+1} - y_k^i \to 0$ quand $i \to +\infty$</u>

   Pour tout i, nous avons

$$\begin{aligned}
0 = f_k(x^k) &\geq \overline{f}_k^{i+1}(x^k) \quad \text{(par C1)}\\
&= \widetilde{f}_k^{i+1}(x^k) \quad \text{(par (6.14))}\\
&\geq \widetilde{f}_k^{i+1}(y_k^{i+1}) \quad \text{(par définition de } y_k^{i+1})\\
&= \widetilde{l}_k^{i+1}(y_k^{i+1}) \quad \text{(par (6.15))}\\
&\geq \widetilde{l}_k^{i}(y_k^{i+1}) \quad \text{(par (6.16))}\\
&= \widetilde{l}_k^{i}(y_k^{i}) + \frac{1}{\epsilon_k} D_h(y_k^{i+1}, y_k^i) \quad \text{(par (6.13))}\\
&\geq \widetilde{l}_k^{i}(y_k^{i}) + \frac{\kappa}{2\epsilon_k}\|y_k^{i+1} - y_k^i\|^2 \quad \text{(par convexité forte de } h \text{ sur } C)\\
&\geq \widetilde{l}_k^{i}(y_k^{i})
\end{aligned}$$

   où $D_h(y, z) = h(y) - h(z) - \langle \nabla h(z), y - z\rangle$. De ces relations, nous avons pour tout i, que

$$\widetilde{l}_k^{i+1}(y_k^{i+1}) \geq \widetilde{l}_k^{i}(y_k^{i}).$$

   Donc, la suite $\{\widetilde{l}_k^i(y_k^i)\}_{i\in\mathbb{N}_0}$ est croissante et bornée supérieurement par 0. Par conséquent $\{\widetilde{l}_k^i(y_k^i)\}_{i\in\mathbb{N}_0}$ est convergente et $y_k^{i+1} - y_k^i \to 0$ quand $i \to +\infty$.

(b) La suite $\{y_k^i\}_{i\in\mathbb{N}_0}$ est bornée

Nous avons, pour y fixé,

$$\begin{aligned}
& f_k(y) + \frac{1}{\epsilon_k}\{h(y) - h(x^k) - \langle \nabla h(x^k), y - x^k\rangle\} \\
& \quad \geq \overline{f}_k^{i+1}(y) + \frac{1}{\epsilon_k}\{h(y) - h(x^k) - \langle \nabla h(x^k), y - x^k\rangle\} \qquad \text{(par C1)} \\
& \quad = \widetilde{f}_k^{i+1}(y) \qquad \text{(par définition)} \\
& \quad \geq \widetilde{l}_k^i(y) \qquad \text{(par (6.26))} \\
& \quad = \widetilde{l}_k^i(y_k^i) + \frac{1}{\epsilon_k}\{h(y) - h(y_k^i) - \langle \nabla h(y_k^i), y - y_k^i\rangle\} \qquad \text{(par (6.13))} \\
& \quad \geq \widetilde{l}_k^i(y_k^i) + \frac{\kappa}{2\epsilon_k}\|y - y_k^i\|^2
\end{aligned}$$

où la dernière inégalité est obtenue par convexité forte de h sur C.

Puisque la suite $\{\widetilde{l}_k^i(y_k^i)\}_{i\in\mathbb{N}_0}$ est convergente, la suite $\{y - y_k^i\}_{i\in\mathbb{N}_0}$ est bornée et donc la suite $\{(y_k^i)\}_{i\in\mathbb{N}_0}$ est aussi bornée.

(c) $f_k(y_k^{i+1}) - \overline{f}_k^{i+1}(y_k^{i+1}) \to 0$

Nous avons successivement par (C2), (C1) et la définition de $s(y_k^{i+1})$,

$$\begin{aligned}
\langle s(y_k^i), y_k^{i+1} - y_k^i\rangle &\leq \overline{f}_k^{i+1}(y_k^{i+1}) - f_k(y_k^i) \\
&\leq f_k(y_k^{i+1}) - f_k(y_k^i) \\
&\leq \langle s(y_k^{i+1}), y_k^{i+1} - y_k^i\rangle.
\end{aligned}$$

Puisque $\{y_k^i\}_{i\in\mathbb{N}_0}$ est bornée, il suit que l'ensemble $\bigcup_i \partial f_k(y_k^i)$ l'est aussi et donc la suite $\{s(y_k^i)\}_{i\in\mathbb{N}_0}$ est bornée. Donc, nous obtenons

$$f_k^{i+1}(y_k^{i+1}) - f_k(y_k^i) \to 0 \quad \text{et} \quad f_k(y_k^{i+1}) - f_k(y_k^i) \to 0$$

et, par conséquent

$$f_k(y_k^{i+1}) - \overline{f}_k^{i+1}(y_k^{i+1}) = f_k(y_k^{i+1}) - f_k(y_k^i) + f_k(y_k^i) - \overline{f}_k^{i+1}(y_k^{i+1}) \to 0.$$

2. $y_k^i \to \overline{y}_k \equiv \arg\min_{y\in C}\{\epsilon_k f_k(y) + h(y) - h(x^k)\langle \nabla h(x^k), y - x^k\rangle\}$

Puisque la suite $\{y_k^i\}_{i\in\mathbb{N}_0}$ est bornée, il reste à prouver que chaque valeur d'adhérence y_k^* de cette suite est égale à $\overline{y}_k$, i.e., que

$$\frac{1}{\epsilon_k}\{\nabla h(x^k) - \nabla h(y_k^*)\} \in \partial(f_k + \delta_C)(y_k^*)$$

ou, par définition de sous-différentielle, nous obtenons de (C1) et (6.9) que

$$\forall y \in C \quad f_k(y) \geq \overline{f}^i_k(y) \geq \overline{f}^i_k(y^i_k) + \frac{1}{\epsilon_k}\langle \nabla h(x^k) - \nabla h(y^i_k), y - y^i_k\rangle$$

i.e.

$$\begin{aligned} \forall y \in C \quad f_k(y) \quad \geq \quad & [\overline{f}^i_k(y) - f_k(y^i_k)] + [f_k(y^i_k) - f_k(y^*_k)] \\ & + f_k(y^*_k) + \frac{1}{\epsilon_k}\langle \nabla h(x^k) - \nabla h(y^i_k), y - y^i_k\rangle. \end{aligned} \tag{6.17}$$

Puisque y^*_k est une valeur d'adhérence de $\{y^i_k\}_{i\in\mathbb{N}_0}$, il existe $\mathbb{K} \subset \mathbb{N}_0$ tel que

$$y^i_k \to y^*_k \text{ pour } i \in \mathbb{K}, i \to +\infty.$$

En prenant la limite (pour $i \in \mathbb{K}$) des deux côtés de (6.17), nous obtenons, pour tout $y \in C$, que

$$\begin{aligned} f_k(y) \quad \geq \quad & \lim_i[\overline{f}^i_k(y^i_k) - f_k(y^i_k)] + \lim_i[f_k(y^i_k) - f_k(y^*_k)] + f_k(y^*_k) \\ & + \frac{1}{\epsilon_k}\lim_i\langle \nabla h(x^k) - \nabla h(y^i_k), y - y^i_k\rangle. \end{aligned}$$

Puisque $\lim_i[\overline{f}^i_k(y^i_k) - f_k(y^i_k)] = 0$ par (1), $\lim_i[f_k(y^i_k) - f_k(y^*_k)] = 0$ parce que f_k est continue, et ∇h est continu en y^*_k, nous déduisons que

$$f_k(y) \geq f_k(y^*_k) + \frac{1}{\epsilon_k}\langle \nabla h(x^k) - \nabla h(y^*_k), y - y^*_k\rangle \quad \forall y \in C.$$

Ceci complète la preuve. □

Théorème 6.2.1 *Supposons que x^k ne soit pas une solution du problème (EP). Alors l'algorithme de pas sérieux s'arrête après un nombre fini d'itérations i_k avec $\overline{f}^{i_k}_k$ une μ-approximation de f_k en x^k et avec $x^{k+1} = y^{i_k}_k$.*

Preuve.
Supposons, pour avoir une contradiction, que le STOP n'ait jamais lieu. Alors

$$f_k(y^i_k) > \mu\overline{f}^i_k(y^i_k) \quad \text{pour tout } i \in \mathbb{N}_0. \tag{6.18}$$

De plus, par le Lemme 6.2.1, $y^i_k \to \overline{y}_k$. Alors, en prenant la limite des deux membres de (6.18), nous obtenons

$$f_k(\overline{y}_k) \geq \mu f_k(\overline{y}_k)$$

parce que f_k est continu sur C et $f_k(y^i_k) - \overline{f}^i_k(y^i_k) \to 0$. D'où, parce que $\mu < 1$, nous déduisons que $f_k(\overline{y}_k) \geq 0$.

D'un autre côté, par définition de $\overline{y}_k$ (voir Lemme 6.2.1), nous avons, pour tout $y \in C$,

$$\epsilon_k f_k(\overline{y}_k) + h(\overline{y}_k) - h(x^k) - \langle \nabla h(x^k), \overline{y}_k - x^k \rangle \leq \epsilon_k f_k(y) + h(y) - h(x^k) - \langle \nabla h(x^k), y - x^k \rangle.$$

Si nous choisissons $y = x^k$ et si nous observons que $f_k(x_k) = 0$, alors l'inégalité devient

$$\epsilon_k f_k(\overline{y}_k) \leq -h(\overline{y}_k) + h(x^k) + \langle \nabla h(x^k), \overline{y}_k - x^k \rangle.$$

Enfin, en utilisant la forte convexité de h cela donne

$$0 \leq \epsilon_k f_k(\overline{y}_k) \leq -\frac{\kappa}{2} \|\overline{y}_k - x^k\|.$$

Par conséquent $\|\overline{y}_k - x^k\| = 0$ et donc $x^k = \overline{y}_k$. Mais cela signifie que x^k est une solution du problème (EP), ce qui contredit les hypothèses du théorème. Donc l'algorithme de pas sérieux s'arrête après un nombre fini d'itérations. $\square$

En incorporant l'algorithme de pas sérieux dans le Pas 1 de l'algorithme général, nous obtenons l'algorithme suivant.

Algorithme 6.2.2 ***Algorithme faisceau pour résoudre le problème*** (EP) :

Soit $x^0 \in C$ un point initial, $\mu \in (0, 1)$ une tolérance et $\{\epsilon_k\}_{k \in I\!N}$ une suite positive. Poser $y_0^0 = x^0$, $k = 0$ et $i = 1$.

Pas 1. *Choisir une fonction convexe linéaire par morceaux $\overline{f}_k^i$ satisfaisant (C1)-(C3) et résoudre*

$$(P_k^i) \quad \min_{y \in C} \{\epsilon \overline{f}_k^i(y) + h(y) - h(x^k) - \langle \nabla h(x^k), y - x^k \rangle\}$$

pour obtenir la solution optimale unique $y_k^i \in C$.

Pas 2. *Si*

$$f_k(y_k^i) \leq \mu \overline{f}_k^i(y_k^i) \tag{6.19}$$

alors poser $x^{k+1} = y_k^i$, $y_{k+1}^0 = x^{k+1}$, $k = k+1$ et $i = 0$.

Pas 3. *$i = i + 1$ et retour au Pas 1.*

Des Théorèmes 6.1.3 et 6.2.1, nous obtenons les résultats de convergence suivants :

Théorème 6.2.2 *Si après qu'un certain k ait été atteint, le critère (6.19) n'est jamais satisfait, alors x^k est une solution du problème (EP).*

Théorème 6.2.3 *Supposons que $\epsilon_k \geq \epsilon \geq 0$ pour tout $k \in I\!N$ et que toutes les hypothèses du Théorème 6.1.2 soient satisfaites, et que la suite $\{x^k\}$ générée par l'algorithme faisceau soit infinie. Alors $\{x^k\}$ converge vers une certaine solution du problème (EP).*

Pour une implémentation pratique, il est nécessaire de donner un critère d'arrêt. Afin de le présenter, nous introduisons la définition de point stationnaire.

Définition 6.2.1 *Soit $\Delta \geq 0$. Un point $x^* \in I\!R^n$ est appelé un point Δ-stationaire du problème (EP) si $x^* \in C$ et si*

$$\exists \gamma \in \partial_\Delta (f_{x^*} + \delta_C)(x^*) \quad \textit{tel que} \quad \|\gamma\| \leq \Delta.$$

Proposition 6.2.1 *Soit y_k^i une solution du problème (P_k^i) et soient*

$$\gamma_k^i = \frac{1}{\epsilon_k}[\nabla h(x^k) - \nabla h(y_k^i)] \quad \textit{et} \quad \delta_k^i := \langle \gamma_k^i, y_k^i - x^k \rangle - \overline{f}_k^i(y_k^i). \tag{6.20}$$

Alors $\quad \delta_k^i \geq 0 \quad$ *et* $\quad \gamma_k^i \in \partial_{\delta_k^i}(f_k + \delta_C)(x^k).$

Preuve.
Par optimalité de y_k^i, nous obtenons que

$$0 \in \epsilon_k \partial(\overline{f}_k^i + \delta_C)(y_k^i) + \nabla h(y_k^i) - \nabla h(x^k)$$

i.e.

$$\gamma_k^i \in \partial(\overline{f}_k^i + \delta_C)(y_k^i).$$

D'où par définition de sous-différentielle et puisque $\overline{f}_k^i \leq f_k$, nous avons, pour tout $x \in C$

$$f_k(x) \geq \overline{f}_k^i(x) \geq \overline{f}_k^i(y_k^i) + \langle \gamma_k^i, x - y_k^i \rangle. \tag{6.21}$$

En particulier pour $x = x^k$, et en remarquant que $f_k(x^k) = 0$, nous déduisons que

$$0 \geq \overline{f}_k^i(y_k^i) + \langle \gamma_k^i, x^k - y_k^i \rangle$$

i.e., $\delta_k^i \geq 0$.

D'un autre côté, par (6.21) et la définition de δ_k^i, nous pouvons écrire pour tout $x \in C$,

$$f_k(x) \geq \overline{f}_k^i(y_k^i) + \langle \gamma_k^i, x - y_k^i \rangle = f_k(x^k) + \langle \gamma_k^i, x - x^k \rangle - \delta_k^i$$

i.e., $\gamma_k^i \in \partial_{\delta_k^i}(f_k + \delta_C)(x^k)$. □

Théorème 6.2.4 *Supposons que $\epsilon_k \geq \epsilon > 0$ pour tout $k \in I\!N$ et que toutes les hypothèses du Théorème 6.1.2 aient lieu. Soit $\{x^k\}$ la suite générée par l'algorithme faisceau.*

1. *Si $\{x^k\}$ est infinie, alors les suites $\{\gamma_k^{i_k}\}_k$ et $\{\delta_k^{i_k}\}_k$ convergent vers zéro.*
2. *Si $\{x^k\}$ est finie avec k le dernier indice, alors les suites $\{\gamma_k^i\}_i$ et $\{\delta_k^i\}_i$ convergent vers zéro.*

Preuve.

1. Puisque $\{x^k\}_k$ est infinie, il suit du Théorème 6.2.3 que $\{x^k\}$ converge vers une certaine solution x^* du problème (EP).
D'un autre côté, nous avons, pour tout k

$$0 \leq \|\gamma_k^{i_k}\| = \left\| \frac{\nabla h(x^k) - \nabla h(y_k^{i_k})}{\epsilon_k} \right\| \leq \frac{L}{\epsilon}\|x^k - y_k^{i_k}\| = \frac{L}{\epsilon}\|x^k - x^{k+1}\|$$

parce que ∇h est Lipschitz continu de constante $L > 0$, $\epsilon_k \geq \epsilon > 0$ et $y_k^{i_k} = x^{k+1}$. Puisque $\|x^{k+1} - x^k\| \to 0$, nous obtenons que la suite $\{\gamma_k^{i_k}\}_k$ converge vers 0. De plus, puisque

$$|\langle \gamma_k^{i_k}, y_k^{i_k} - x^k \rangle| \leq \|\gamma_k^{i_k}\|\|y_k^{i_k} - x^k\| = \|\gamma_k^{i_k}\|\|x^{k+1} - x^k\|$$

nous obtenons aussi que $\langle \gamma_k^{i_k}, y_k^{i_k} - x^k \rangle \to 0$ quand $k \to +\infty$. Enfin, par définition de x^{k+1} et (6.19), nous avons

$$\frac{1}{\mu} f_k(x^{k+1}) \leq \overline{f}_k^{i_k}(x^{k+1}) \leq f_k(x^{k+1}). \tag{6.22}$$

Puisque $f_k(x^{k+1}) = f(x^k, x^{k+1}) \to f(x^*, x^*) = 0$ par continuité de f_k, (6.22) implique que $\overline{f}_k^{i_k}(x^{k+1}) \to 0$. Par conséquent, nous obtenons que $\delta_k^i \to 0$ quand $k \to +\infty$.

2. Soit k le dernier indice de la suite $\{x^k\}$. Alors x^k est une solution du problème (EP) par le Théorème 6.2.2 et $\{y_k^i\}_i$ converge vers $\overline{y}_k$ quand $i \to +\infty$ par le Lemme 6.2.1. D'où $x^k = \overline{y}_k$ et $\|x^k - y_k^i\| \to 0$ quand $i \to +\infty$. Mais cela signifie que $\{\gamma_k^i\}_i$ converge vers zéro. De plus, par le Lemme 6.2.1, pour $\{\gamma_k^i\}_i$, nous avons $f_k(y_k^i) - \overline{f}_k^i(y_k^i) \to 0$ et donc $\overline{f}_k^i(y_k^i) = \overline{f}_k^i(y_k^i) - f_k(y_k^i) + f_k(y_k^i) \to 0$ parce que f_k est continue et $f_k(y_k^i) = f(x^k, y_k^i) \to f(x^k, x^k) = 0$. Par conséquent $\delta_k^i \to 0$ quand $i \to +\infty$. □

Grâce à la Proposition 6.2.1 et au Théorème 6.2.4, nous pouvons introduire un critère d'arrêt dans l'algorithme faisceau juste après le Pas 1 comme suit :

Calculer γ_k^i et δ_k^i en utilisant (6.20). Si $\|\gamma_k^i\| \leq \Delta$ et $\delta_k^i \leq \Delta$, alors STOP ; x^k est un point Δ-stationnaire du problème (EP). Autrement, retour au Pas 2 de l'algorithme faisceau.

Pas 2. *Si*

$$\varphi(x^k) - \varphi(y_k^i) \geq \mu[\varphi(x^k) - \theta_k^i(y_k^i)] + (1-\mu)\langle F(x^k), y_k^i - x^k \rangle \tag{6.24}$$

alors poser $x^{k+1} = y_k^i$, $y_{k+1}^0 = x^{k+1}$, $k = k+1$ *et poser* $i = 0$.

Pas 3. $i = i + 1$ *et retour au Pas 1.*

Théorème 6.3.1 *Supposons que la suite* $\{\epsilon_k\}$ *soit décroissante et satisfasse* $0 < \epsilon \leq \epsilon_k$ *pour tout* k. *Si* F *est* φ*-co-coercif sur* C *avec* $\gamma > \frac{\epsilon_0}{2\kappa\mu^2}$, *et si la suite* $\{x^k\}$ *générée par l'algorithme faisceau pour résoudre* $(GVIP)$ *est infinie, alors la suite* $\{x^k\}$ *converge vers une certaine solution du problème* $(GVIP)$.

Preuve.
Par le Théorème 6.2.3 et la Proposition 6.1.2, nous devons seulement prouver que si $\gamma > \frac{\epsilon_0}{2\kappa\mu^2}$ alors il existe $\tau > 0$ tel que $\epsilon_0 < \frac{\kappa\mu}{\tau}$ et $\mu \geq \frac{1}{2\tau\gamma}$. Puisque $\epsilon_0 < 2\kappa\mu^2\gamma$ il est suffisant de poser $\tau = \frac{1}{2\mu\gamma} > 0$ pour obtenir les deux inégalités. □

6.3.2 Problème d'inéquation variationnelle multivoque $(MVIP)$

Comme seconde application, nous appliquons l'algorithme général au problème $(MVIP)$. Ce problème correspond au problème (EP) avec la fonction f définie, pour tout $x, y \in C$, par $f(x,y) = \sup_{\xi \in F(x)} \langle \xi, y - x \rangle$ où $F : C \to 2^{\mathbb{R}^n}$ est un opérateur multivoque continu à valeurs compactes et f est continue sur $C \times C$. A l'itération k, nous considérons la fonction approximante $\overline{f}_k(y) = \langle \xi^k, y - x^k \rangle$ avec $\xi^k \in F(x^k)$. Ici, nous supposons qu'au moins un élément de $F(x)$ est disponible pour chaque $x \in C$. Quand h est la norme carré, le sous-problème $(\overline{P}_k)$ devient

$$\min_{y \in C} \left\{ \epsilon_k \langle \xi^k, y - x^k \rangle + \frac{1}{2} \|y - x^k\|^2 \right\}. \tag{6.25}$$

Nous observons que les conditions d'optimalité associées à (6.25) sont

$$\langle \epsilon_k \xi^k + y^k - x^k, y - y^k \rangle \geq 0 \quad \forall y \in C \tag{6.26}$$

où y^k est une solution de $(\overline{P}_k)$. En d'autres mots, y^k est la projection orthogonale du vecteur $x^k - \epsilon_k \xi^k$ sur C. Ce problème est un problème de programmation quadratique convexe particulier dont la solution peut être trouvée explicitement quand C a une structure spéciale comme une boîte, une boule, ... Sans perdre de généralité, nous pouvons supposer que $y^k \neq x^k$. En effet, si $y^k = x^k$, alors il est facile de voir que x^k est une solution de $(MVIP)$.

Notre but est tout d'abord de trouver des conditions pour assurer que la fonction $\overline{f}_k$ définie plus haut soit une μ-approximation de f_k en x^k et alors les appliquer au Théorème 6.1.3 pour obtenir la convergence de la suite $\{x^k\}$. Dans ce but, nous introduisons les définitions suivantes.

Définition 6.3.1 *Soit C un sous-ensemble convexe fermé non vide de $I\!R^n$ et soit $F : C \to 2^{I\!R^n}$.*

1. *F est fortement monotone sur C si $\exists \alpha > 0$ tel que $\forall x, y \in C$, $\forall \xi_1 \in F(x)$, $\xi_2 \in F(y)$, nous avons*
$$\langle \xi_1 - \xi_2, x - y \rangle \geq \alpha \|x - y\|^2.$$
2. *F est Lipschitz continu sur C si $\exists L > 0$ tel que $\forall x, y \in C$, nous avons*
$$g(x, y) \leq L\|x - y\|$$
où
$$g(x, y) := \sup_{\xi_1 \in F(x)} \inf_{\xi_2 \in F(y)} \|\xi_1 - \xi_2\|^2. \tag{6.27}$$
3. *F est co-coercif sur C si $\exists \gamma > 0$ tel que $\forall x, y \in C$, $\forall \xi_1 \in F(x)$, $\forall \xi_2 \in F(y)$, nous avons*
$$\langle \xi_1 - \xi_2, x - y \rangle \geq \gamma g(x, y).$$

Dans la proposition suivante, nous présentons la propriété principale de la fonction $\overline{f}_k$.

Proposition 6.3.1 *Supposons que F soit co-coercif sur C de constante $\gamma > 0$. Soit $\mu \in (0, 1)$ et $x^k \in C$. Si $\epsilon_k \leq 4\gamma(1 - \mu)$, alors la fonction $\overline{f}_k(y) = \langle \xi_k, y - x^k \rangle$ avec $\xi_k \in F(x^k)$ est une μ-approximation de f_k en x^k, i.e., $\overline{f}_k \leq f_k$ et $\overline{f}_k(y^k) \leq \mu f_k(y^k)$ où y^k est une solution du problème (P_k).*

Preuve.
Soit $\mu \in (0, 1)$ et ξ_k, $\xi \in F(x^k)$. Par (6.26), avec $y = x^k$, nous déduisons que
$$\epsilon_k \langle \xi_k, y^k - x^k \rangle \leq -\|x^k - y^k\|^2 < 0. \tag{6.28}$$

En utilisant successivement la co-coercivité de F, la définition de g en (6.27) et l'inégalité de Cauchy-Schwarz, nous avons pour chaque $\eta \in F(y^k)$ et pour n'importe quel $\nu > 0$
$$\begin{aligned}\langle \xi - \xi_k, y^k - x^k \rangle &= \langle \xi - \eta, y^k - x^k \rangle + \langle \eta - \xi_k, y^k - x^k \rangle \\ &\leq -\gamma g(x^k, y^k) + \|\eta - \xi_k\| \|y^k - x^k\| \\ &\leq -\gamma g(x^k, y^k) + (1/2\nu)\|\eta - \xi_k\|^2 + \nu/2\|y^k - x^k\|^2.\end{aligned} \tag{6.29}$$

En prenant l'infimum sur $\eta \in F(y^k)$ et en utilisant (6.28), nous obtenons

$$\begin{aligned}\langle \xi - \xi_k, y^k - x^k\rangle &\leq -\gamma g(x^k, y^k) + \frac{1}{2\nu}\inf_{\eta \in F(y^k)} \|\eta - \xi_k\|^2 + \frac{\nu}{2}\|y^k - x^k\|^2 \\ &\leq (\frac{1}{2\nu} - \gamma) g(x^k, y^k) - \frac{\nu \epsilon_k}{2}\langle \xi_k, y^k - x^k\rangle\end{aligned}$$

pour tout $\nu > 0$. Choisissons $\nu = 1/(2\gamma)$, nous pouvons écrire

$$\langle \xi, y^k - x^k\rangle \leq (1 - \frac{\epsilon_k}{4\gamma})\langle \xi_k, y^k - x^k\rangle.$$

Enfin, en prenant le supremum sur $\xi \in F(x^k)$, et en utilisant la condition $\epsilon_k \leq 4\gamma(1-\mu)$, nous déduisons cette thèse. □

Puisque $\overline{f}_k$ est une μ-approximation de f_k en x^k pour une valeur adaptée de ϵ_k, en utilisant cette fonction approximante, l'algorithme général devient :

Soit $x^k \in C$ donné et $\epsilon_k > 0$, choisir $\xi_k \in F(x^k)$ et résoudre le problème

$$\min_{y \in C}\{\epsilon_k \langle \xi_k, y - x^k\rangle + \frac{1}{2}\|y - x^k\|^2\}$$

pour obtenir x^{k+1}.

En particulier, quand F est co-coercif sur C, les hypothèses (1) et (2) du Théorème 6.1.2 sont satisfaites.

Proposition 6.3.2 *Soit $f(x,y) = \sup_{\xi \in F(x)}\langle \xi, y - x\rangle$ et g défini par (6.27). Alors*

1. *pour chaque $x, y, z \in C$ et pour n'importe quel $\nu > 0$,*

$$f(x,z) - f(y,z) - f(x,y) \leq \frac{1}{2\nu} g(x,y) + \frac{\nu}{2}\|z - y\|^2 \,;$$

2. *si F est co-coercif sur C de constante γ, alors pour chaque $x, y \in C$,*

$$f(x,y) \geq 0 \Rightarrow f(y,x) \leq -\gamma g(x,y).$$

Enfin, pour la suite $\{x^k\}$ générée par cet algorithme, nous obtenons le théorème de convergence suivant.

Théorème 6.3.2 *Supposons que F soit co-coercif sur C de constante $\gamma > 0$. Soit $\{\epsilon_k\}$ une suite décroissante bornée par 0. Si $\epsilon_k < 4(2-\sqrt{3})\gamma$ pour tout k, alors la suite $\{x^k\}$ converge vers une certaine solution x^* du problème $(MVIP)$.*

Preuve.
Puisque $\kappa = 1$, par les Propositions 6.3.1 et 6.3.2, et par le Théorème 6.1.3, nous devons seulement prouver qu'il existe $\mu \in (0,1)$, et $\nu > 0$ tel que

$$\epsilon_k \leq 4\gamma(1-\mu), \quad \epsilon_k < \frac{\mu}{\nu}, \quad \frac{1}{2\nu\gamma} \leq \mu.$$

Choisissant le plus petit ν possible, nous obtenons $\nu = 1/(2\mu\gamma)$. Alors les conditions précédentes deviennent :

$$\epsilon_k \leq 4\gamma(1-\mu) \quad \text{et} \quad \epsilon_k < 2\mu^2\gamma. \tag{6.30}$$

Il est facile de voir que le maximum de la fonction $r(\mu) = \min\{4\gamma(1-\mu), 2\mu^2\gamma\}$ est atteint en $\mu = \sqrt{3}-1$ et a $4(2-\sqrt{3})\gamma$ comme valeur optimale. Donc, les conditions (6.30) sont satisfaites avec cette valeur de μ si $\epsilon_k < 4(2-\sqrt{3})\gamma$. □

Quand F est univoque, la fonction approximante $\overline{f}_k$ coïncide avec f_k. Dans ce cas, $\mu = 1$ et la Proposition 6.3.1 ne doit pas être considérée. Cela signifie que seulement la seconde inégalité dans (6.30) doit être retenue, i.e.,

$$\epsilon_k < 2\gamma.$$

Un cas particulier intéressant est quand F est fortement monotone (de constante $\alpha > 0$) et Lipschitz continu (de constante $L > 0$) sur C. Alors, pour tout $x, y \in C$,

$$g(x,y) \leq L^2\|x-y\|^2.$$

D'où F étant fortement monotone, nous avons pour tout $x, y \in C$ et $\xi \in F(x)$, $\eta \in F(y)$, que

$$\langle \xi - \eta, x - y\rangle \geq \alpha\|x-y\|^2 \geq \frac{\alpha}{L^2} g(x,y).$$

Mais cela signifie que F est co-coercif sur C de constante $\gamma = \alpha/L^2$. Alors, le Théorème 6.3.2 devient :

Théorème 6.3.3 *Supposons que F soit fortement monotone (de constante $\alpha > 0$) et Lipschitz continu (de constante $L > 0$) sur C. Soit $\{\epsilon_k\}$ une suite décroissante bornée par 0. Si $\epsilon_k < 4(2-\sqrt{3})\frac{\alpha}{L^2}$ pour tout k, alors la suite $\{x^k\}$ converge vers l'unique solution x^* du problème $(MVIP)$. Quand F est univoque, la même propriété a lieu mais avec $\epsilon_k < \frac{2\alpha}{L^2}$ pour tout k.*

Conclusion

Dans le premier chapitre, nous avons défini le problème d'équilibre ainsi que quelques problèmes apparentés.

Au second chapitre, nous avons décrit une méthode de relaxation combinée pour résoudre un problème d'équilibre.

Dans les trois chapitres qui suivent, nous avons ensuite décrit des méthodes de relaxations combinées pour résoudre les problèmes d'inéquations variationnelles généralisés et multivoques ainsi que les inclusions multivoques.

Dans le dernier chapitre, nous avons décrit une méthode faisceau pour résoudre le problème d'équilibre que nous avons appliquée aux problèmes d'inéquations variationnelles généralisés et multivoques.

Pour chacune de ces méthodes décrites, nous en avons étudié la convergence.

Annexe A

Définitions et divers résultats

Définition A.0.1 *Soient W et V des ensembles convexes dans $I\!R^n$, $W \subseteq V$ et soit $\varphi : V \to I\!R \bigcup \{+\infty\}$ une fonction. La fonction φ est dite*

1. *propre s'il existe un point $x \in V$ tel que $\varphi(x) < +\infty$;*
2. *sous-différentiable en $x \in V$ si $\partial\varphi(x) \neq \emptyset$;*
3. *semicontinue supérieurement (inférieurement) (u.s.c. (l.s.c.)) sur W, si pour chaque suite $\{x^k\} \to \overline{x}$, $x^k \in W$ nous avons*

$$\limsup_{k\to+\infty} \varphi(x^k) \leq \varphi(\overline{x}) \quad \left(\liminf_{k\to+\infty} \varphi(x^k) \geq \varphi(\overline{x})\right).$$

Définition A.0.2 *Soient W et V deux ensembles convexes dans $I\!R^n$ tels que $W \subseteq V$. Soit $P : V \to \Pi(I\!R^n)$ un opérateur multivoque. L'opérateur P est dit : x-hemicontinu sur W, si pour tout $x \in W$, $y \in W$ et $\alpha \in [0,1]$, l'opérateur $\alpha \to \langle Q(x+\alpha z), w\rangle$ avec $z = y - x$ est u.s.c. en 0^+.*

Proposition A.0.1 *Supposons que x^* soit un minimum d'une fonction $\varphi : I\!R^n \to I\!R \bigcup \{+\infty\}$ sur $I\!R^n$. Si φ est sous-différentiable en x^*, alors $0 \in \partial\varphi(x^*)$.*

Pour un ensemble convexe W, nous définissons une fonction indicatrice $\psi_W : \mathbb{R}^n \to \mathbb{R} \bigcup \{+\infty\}$ comme suit :

$$\psi_W(y) = \begin{cases} 0 & \text{si } y \in W, \\ +\infty & \text{si } y \notin W. \end{cases}$$

Nous avons

$$\partial\psi_W(y) = \begin{cases} N(W,y) & \text{si } y \in W, \\ +\emptyset & \text{si } y \notin W. \end{cases}$$

Annexe B

Problème d'équilibre et problème d'inéquation variationnelle généralisé

Corollaire B.0.1 *Soit $F : C \to I\!R^n$ un opérateur continu.*

1. *Si F est strictement monotone, alors $(GVIP)$ a au plus une solution.*
2. *Si F est fortement monotone, alors $(GVIP)$ a une solution unique.*

Proposition B.0.1 *Soient $F : C \to I\!R^n$ un opérateur continu, $\varphi : I\!R^n \to I\!R \bigcup \{+\infty\}$ une fonction l.s.c., propre, convexe et sous-différentiable.*

1. *Le problème $(GVIP)$ est équivalent au problème de trouver $x^* \in C$ tel que*

$$\exists t^* \in \partial\varphi(x^*), \quad \langle F(x^*) + t^*, x - x^* \rangle \geq 0 \quad \forall x \in C. \tag{B.1}$$

2. *Si φ est strictement convexe, alors $(GVIP)$ a au moins une solution.*
3. *Si φ est fortement convexe, alors $(GVIP)$ a une solution unique.*

Théorème B.0.1 *Soit W un ensemble convexe ouvert dans $I\!R^n$ tel que $C \subseteq W$. Soit $f : W \times W \to I\!R \bigcup \{+\infty\}$ une bifonction d'équilibre et soit F défini par $F(x) = \partial f_y(x, y)|_{y=x}$ un opérateur. Soient C^0 et $C^0_{(d)}$ les ensembles de solutions de (EP) et (DEP) respectivement et C^* et C^d les ensembles de solutions de (GVIP) et (DGVIP) respectivement.*

1. *Si $f(x, \cdot)$ est quasiconvexe et régulière pour chaque $x \in W$, alors*

$$C^0 \subseteq C^*, \quad C^0_{(d)} \subseteq C^d.$$

2. *Si $f(x, \cdot)$ est pseudoconvexe et l.s.c. pour chaque $x \in W$, alors*

$$C^0 = C^*, \quad C^0_{(d)} \subseteq C^d.$$

Annexe C

Problème d'inéquation variationnelle multivoque

Proposition C.0.1 *1. L'ensemble C^d est convexe et fermé.*

2. Si F est hemicontinu et a des valeurs convexes non vides et compactes, alors $C^d \subseteq C^*$.

3. Si F est pseudomonotone, alors $C^* = C^d$.

Lemme C.0.1 *Soient $\{p^i\}$ et $\{q^i\}$ deux suites dans $I\!R^n$ telles que*

$$\|q^i\| \leq C \leq +\infty, \qquad p^{i+1} = Nrconv\{p^i, q^{i+1}\} \quad , \quad i = 0, 1, \dots \qquad p^0 = q^0 \tag{C.1}$$

Alors

$$p^i \in conv\{q^0, \dots, q^i\}, \quad i = 0, 1, \dots$$

Lemme C.0.2 *Soient $\{p^i\}$ et $\{q^i\}$ deux suites dans $I\!R^n$ telles que (C.1) ait lieu et*

$$\langle q^{i+1}, p^i \rangle \leq \theta \|p^i\|^2, \quad i = 0, 1, \dots, \quad \theta \in (0, 1)$$

Alors

$$\|p^i\| \leq \frac{C}{(1-\theta)\sqrt{i+1}} \quad pour\ i = 0, 1, \dots$$

Bibliographie

[1] Gol'shtein E.G. and Tret'yakov N.V., *Augmented Lagrange Functions*, Nauka, Moscow, 1989.

[2] Hiriart-Urruty J.B. and Lemaréchal C., *Convex Analysis and Minimization Algorithms*, Springer-Verlag, Berlin, 1996.

[3] Konnov I.V., *The application of a linearization method to solving nonsmooth equilibrium problems*, Russian Mathematics, Vol. 40, pp. 54-62, 1996.

[4] Konnov I.V., *Combined Relaxation Methods for Variational Inequalities*, Springer-Verlag, Berlin, 2001.

[5] Mastroeni G., *On Auxiliary Principle for Equilibrium Problems.* In : P. Daniele, F. Giannessi and A. Maugeri, (eds.), Equilibrium Problems and Variational Models. Kluwer Academic Publishers, Dordrecht, pp. 289-298, 2003.

[6] Nguyen T.T.V., Strodiot J.J. and Nguyen V.H., *A bundle method for solving equilibrium problems*, Mathematical Programming, 2007.

[7] Strodiot J.J., *An introduction to nonsmooth optimization*, Cours donné à l'université des Sciences Naturelles, Ho Chi Minh Ville, Vietnam, 2005.

[8] Zhu D. and Marcotte P., *Co-coercivity and its role in the convergence of iterative schemes for solving variational inequalities*, SIAM J. Optim. 6, 714-726, 1996.